

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 11:04:50 ; Search time 126 Seconds  
(without alignments)  
1080.856 Million cell upd

Title: US-10-803-278-4

Perfect score: 2501

Sequence: 1 MLKFQEAACVSGSTAISTY.....LGYHGDCNLISLDEYWKNEK 482

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq 29Jan04:\*

```
1: _ geneseqp1980s:*
```

2: geneseqp1990s:\*

3: geneseqp2000s: \*

4: **geneseqp2001s:\***

5: genesetp2002s:\*

6: genesetp2003as:\*

7: genesetm2003bs:\*

8: geneseqp2004s:\*

.....

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query		Length	DB	ID	Description
	Score	Match				
1	2408	96.3	470	7	ADCS1184	Adc51184 Human cel
2	2408	96.3	645	4	AAU03546	Aau03546 Human pro
3	2408	96.3	645	5	AAE19147	Aae19147 Human kin
4	2408	96.3	645	7	ADCS1186	Adc51186 Human cel
5	2405	96.2	645	5	ABG31081	Abg31081 Human nov
6	2182	87.2	609	7	ADC99091	Adc99091 Human kpp
7	1799.5	72.0	540	7	ADC99090	Adc99090 Human kpp
8	1652	66.1	497	5	ABP68951	Abp68951 Human pol
9	1491	59.6	467	5	ABB72239	Abb72239 Rat prote
10	916	36.6	357	7	ADCS1182	Adc51182 Human cel
11	615	24.6	841	7	ADB37605	Adb37605 Neural th
12	613	24.5	841	7	ADB37570	Adb37570 Human th
13	598.5	23.9	489	6	ABU54632	Abu54632 Human nov
14	598	23.9	774	6	AAO26613	Aao26613 Serine/th
15	598	23.9	774	6	ABP71711	Abp71711 NEK-like
16	598	23.9	774	6	AAO16441	Aao16441 Human ser
17	598	23.9	774	6	ABP97691	Abp97691 Amino aci
18	593.5	23.7	425	6	ABU54633	Abu54633 Human nov
19	592	23.7	506	6	ABU54631	Abu54631 Human nov
20	590.5	23.6	489	7	ADC99067	Adc99067 Human kpp
21	587.5	23.5	507	6	ABP97688	Abp97688 Amino aci
22	585.5	23.4	654	5	AAU77929	Aau77929 Amino aci
23	585	23.4	506	4	AAW78344	Aaw78344 Human pro
24	585	23.4	506	5	ABP60668	Abp60668 Human ser
25	585	23.4	546	6	ABP96072	Abp96072 Human pro

Abp71712	NEK-like
Abp97690	Amino acid
Aae24136	Human actin
Abp72000	Human serum
Abu77928	Amino acid
Aku39211	Human polio
Aku07102	Human non-actin
Aku08113	Human kinase
Abu39210	Human polio
Aao15440	Human protein
Aao03545	Human protein
Abu79228	Human protein
Nbb97324	Novel human
Aku54635	Human NOV
Abp71710	Human NOV
Abu54634	Human NOV
Adc51180	Human NOV
Afu68778	Amino acid
Abp71713	NEK-like
Abp97689	Amino acid

## ALIGNMENTS

## RESULT 1

ADC51184  
ID ADC51184 standard; protein: 470 AA.

ADC51184;

DT 18-DEC-2003 (first entry)

DE Human cell-cycle related protein, SEQ ID 10.

Human; cytostatic; cell-cycle related protein; nuclear export;  
KW  
nuclear-cytoplasm transport; cytotoxic; cell-cycle control;  
KW  
immunological disease; neurological disease; cancer.  
KW

OS Homo sapiens.

AA  
PN JP2003144168-A.

XX  
PD  
20-MAY-2003

14-NOV-2001: 2001JP-00349158

XX  
PR 14-NOV-2001: 2001.TP-00349158

XX  
PA (KOKU-) KOKURITSU YOBO EISEI KENKYUSHO.  
PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.

WPI; 2003-818166/77.  
N-PSDB; ADC51183.

Novel DNA or RNA coding a cell-cycle related protein which has nuclear export function, useful for screening substance that prevent or treat cell cycle abnormality diseases e.g. immunological disease.

PS Disclosure: SEQ ID NO 10: 41bp; Japanese;   
XX

The present invention relates to novel cell-cycle related protein such as NIMA (Never-in Mitosis, gene A)-related protein kinase of Nek 9 and coding sequences such as a cell-cycle related protein (ADC51176) having nuclear export function; cell-cycle related protein (ADC51178) having nuclear-cytoplasm transport function; cell-cycle related protein (ADC51180) having nuclear export function, cytotoxic function and transfer function in the nucleus; and/or cell-cycle related protein (ADC51182) having transfer function and cytotoxic function in nucleus, nuclear export function and nuclear cytoplasm transport function. The coding sequences for these proteins are given in ADC51175, ADC51177, ADC51179 and ADC51181. The sequences of the invention are useful for screening a substance which promotes or suppresses the transfer function

CC or cytotoxic function in the nucleus, nuclear-cytoplasm transport  
CC function and/or a cell-cycle control function. The sequences are also  
CC useful for treating or diagnosing an immunological disease, neurological  
CC disease or cancer. The present sequence was used to illustrate the  
CC invention.  
XX  
SQ Sequence 470 AA;

Query Match		96.3%; Score 2408; DB 7; Length 470;
Best Local Similarity		100.0%; Pred. No. 4.1e-177;
Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	MLKFOBAKCVSGSTAISTYPTKTLIARRVYLOQKLGSGFGTVLVSDKKAKRGELKVL 60
DB	1	MLKFOBAKCVSGSTAISTYPTKTLIARRVYLOQKLGSGFGTVLVSDKKAKRGELKVL 60
QY	61	KEISVGEINPNETVQANLEAQLLSKLDHPAIVKPHASFVEQDNFCIIIEYCEGRDLDDKI 120
DB	61	KEISVGEINPNETVQANLEAQLLSKLDHPAIVKPHASFVEQDNFCIIIEYCEGRDLDDKI 120
QY	121	QBYKQAGKIPFENQIIEWFIQILLGVDMHERRILHRDLKSNVFLKNNLLKIGDFGVS 180
DB	121	QBYKQAGKIPFENQIIEWFIQILLGVDMHERRILHRDLKSNVFLKNNLLKIGDFGVS 180
QY	181	LLMGSCDLATTLTGTPHYMSPEALKHQGYDTKSDIWSLACILYEMCMNHAFAGSNFLSI 240
DB	181	LLMGSCDLATTLTGTPHYMSPEALKHQGYDTKSDIWSLACILYEMCMNHAFAGSNFLSI 240
QY	241	VLKIVEGDTSPSLPERYPKELNAMESMLNKNPSLRPSAIEILKIPYDEQQLNLCRYSE 300
DB	241	VLKIVEGDTSPSLPERYPKELNAMESMLNKNPSLRPSAIEILKIPYDEQQLNLCRYSE 300
QY	301	MTLEDKNLDCQKEAAHIINAMQRIHLQTLRALSEVQKMTPRERMLRKLQAADKARKL 360
DB	301	MTLEDKNLDCQKEAAHIINAMQRIHLQTLRALSEVQKMTPRERMLRKLQAADKARKL 360
QY	361	KKIVEEYKENSCKRMQELRSRNFQQLSVDLVHEKTHLKGMEKEQEPGRJLSCSPQDEDE 420
DB	361	KKIVEEYKENSCKRMQELRSRNFQQLSVDLVHEKTHLKGMEKEQEPGRJLSCSPQDEDE 420

RESULT 2  
AAU03546  
ID AAU03546 standard; protein; 645 AA.  
AC AAU03546;  
XX  
DT 12-SEP-2001 (first entry)  
XX  
DE Human protein kinase #46.  
XX  
KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;  
KW metabolic disorder; immune related disease; neurological disorder;  
KW neurodegenerative disorder; inflammatory disorder; infectious disease;  
KW reproductive disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200138503-A2.  
XX  
PD 31-MAY-2001.  
XX  
PF 22-NOV-2000; 2000WO-US032085.  
XX  
PR 24-NOV-1999; 99US-0167482P.  
XX  
PA (SUGB-) SUGEN INC.  
XX  
PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;

PI Flanagan P, Clary D;  
XX  
DR WPI; 2001-343950/36.  
DR N-PSDB; AAS06746.  
XX  
PT Nucleic acids encoding human kinase polypeptides, useful for preventing  
PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and  
PT neuronal-associated diseases, and microbial infections.  
XX  
PS Claim 7; Fig 2; 433pp; English.

XX AAU03501-AAU03557 represent novel human protein kinases #1-57. The novel  
CC protein kinases have been identified as members of the tyrosine or  
CC serine/threonine kinase (PTK and STK) families. The polynucleotides  
CC encoding protein kinases and the polypeptides may be used in the  
CC prevention, diagnosis and treatment of diseases associated with  
CC inappropriate kinase expression. For example, they may be used to treat  
CC cancers (especially cancers of haematopoietic origin), cardiovascular  
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),  
CC immune related diseases (e.g. rheumatoid arthritis), neurological  
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.  
CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious  
CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).  
CC Additionally, polynucleotides encoding protein kinases may be used for  
CC gene therapy and as DNA probes in diagnostic assays. The protein kinase  
CC polypeptides may be used as antigens in the production of antibodies  
CC against the protein kinases and in assays to identify modulators of  
CC protein kinase expression and activity  
XX  
SQ Sequence 645 AA;

Query Match		96.3%; Score 2408; DB 4; Length 645;
Best Local Similarity		100.0%; Pred. No. 6.2e-177;
Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1	MLKFOBAKCVSGSTAISTYPTKTLIARRVYLOQKLGSGFGTVLVSDKKAKRGELKVL 60
DB	1	MLKFOBAKCVSGSTAISTYPTKTLIARRVYLOQKLGSGFGTVLVSDKKAKRGELKVL 60
QY	61	KEISVGEINPNETVQANLEAQLLSKLDHPAIVKPHASFVEQDNFCIIIEYCEGRDLDDKI 120
DB	61	KEISVGEINPNETVQANLEAQLLSKLDHPAIVKPHASFVEQDNFCIIIEYCEGRDLDDKI 120
QY	121	QBYKQAGKIPFENQIIEWFIQILLGVDMHERRILHRDLKSNVFLKNNLLKIGDFGVS 180
DB	121	QBYKQAGKIPFENQIIEWFIQILLGVDMHERRILHRDLKSNVFLKNNLLKIGDFGVS 180
QY	181	LLMGSCDLATTLTGTPHYMSPEALKHQGYDTKSDIWSLACILYEMCMNHAFAGSNFLSI 240
DB	181	LLMGSCDLATTLTGTPHYMSPEALKHQGYDTKSDIWSLACILYEMCMNHAFAGSNFLSI 240
QY	241	VLKIVEGDTSPSLPERYPKELNAMESMLNKNPSLRPSAIEILKIPYDEQQLNLCRYSE 300
DB	241	VLKIVEGDTSPSLPERYPKELNAMESMLNKNPSLRPSAIEILKIPYDEQQLNLCRYSE 300
QY	301	MTLEDKNLDCQKEAAHIINAMQRIHLQTLRALSEVQKMTPRERMLRKLQAADKARKL 360
DB	301	MTLEDKNLDCQKEAAHIINAMQRIHLQTLRALSEVQKMTPRERMLRKLQAADKARKL 360
QY	361	KKIVEEYKENSCKRMQELRSRNFQQLSVDLVHEKTHLKGMEKEQEPGRJLSCSPQDEDE 420
DB	361	KKIVEEYKENSCKRMQELRSRNFQQLSVDLVHEKTHLKGMEKEQEPGRJLSCSPQDEDE 420
QY	421	ERWQGREESDEPTLENLPESQIPSPMDLHELESIVEDATSDLGYYH 466
DB	421	ERWQGREESDEPTLENLPESQIPSPMDLHELESIVEDATSDLGYYH 466

RESULT 3  
AAE19147  
ID AAE19147 standard; protein; 645 AA.  
XX  
AC AAE19147;

21-MAY-2002 (first entry)  
Human kinase polypeptide (PKIN-5) .  
Human; kinase polypeptide; PKIN-5; gene therapy; Addison's disease; leukemia; immune disorder; lymphoma; melanoma; developmental disorder; acquired immunodeficiency syndrome; AIDS; allergy; anaemia; hypertension; asthma; Crohn's disease; rheumatoid arthritis; bursitis; atherosclerosis; cirrhosis; hepatitis; psoriasis; Cushing's syndrome; cytostatic; cancer; cholestasis; cardiast; cardiovascular disorder; Niemann-Pick's disease; lipid disorder; fatty liver; Gaucher's disease; myocardial infarction; drug screening; transgenic animal; antiinflammatory; hepatotropic; hypotensive; anti-HIV; enzyme.  
Homo sapiens.  
Key Location/Qualifiers  
Domain 29..287  
FT FT /note= "Eukaryotic protein kinase domain"  
Domain 108..121  
FT FT /note= "Tyrosine kinase catalytic domain"  
Domain 148..166  
FT FT /note= "Tyrosine kinase catalytic domain"  
Domain 256..278  
FT FT /note= "Tyrosine kinase catalytic domain"  
WO200208399-A2.  
31-JAN-2002.  
20-JUL-2001; 2001WO-US023092.  
21-JUL-2000; 2000US-0220038P.  
28-JUL-2000; 2000US-0222112P.  
04-AUG-2000; 2000US-0222831P.  
11-AUG-2000; 2000US-0224729P.  
(INCY-) INCYTE GENOMICS INC.  
(THOR-) THORNTON M.  
Thornton M, Yue H, Khan FA, Gururajan R, Hafalia AJA, Wallia NK, Patterson C, Ramkumar J, Gandhi AR, Policky JL, Baughn MR, Tribolety CM, Bandman O, Nguyen DB, Lu Y, Burford N, Lal P, Ding L, Yao MG, Elliott VS, Recipon SA, Kearney L, Lu DAM, Greenwald SR, Tang YT, Xu Y, Walsh RT, Gietzen KJ, Yang J, Hillman JL;  
WPI: 2002-206083/26.  
N-PSDB; AAD30552.  
New human kinase polypeptide, useful in diagnosis, prevention and treatment of cancer, immune disorder, growth and developmental disorder, cardiovascular disorder and lipid disorder.  
Claim 1; Page 138-140; 196pp; English.  
The present invention relates to an isolated human kinase polypeptide (PKIN) or a biologically active/immunogenic fragment of PKIN. PKIN is useful for diagnosing, treating and preventing cancer (e.g., leukaemia, lymphoma, melanoma), an immune disorder (e.g., acquired immunodeficiency syndrome (AIDS), Addison's disease, allergy, anaemia, asthma, Crohn's disease, rheumatoid arthritis), a growth and developmental disorder (e.g. bursitis, cirrhosis, hepatitis, psoriasis, Cushing's syndrome), a cardiovascular disorder (e.g., atherosclerosis, hypertension, myocardial infarction), and a lipid disorder (e.g., fatty liver, cholestasis, Gaucher's disease, Niemann-Pick's disease). PKIN is useful in a number of drug screening techniques and to analyse the proteome of a tissue or cell type. PKIN is useful for creating knockin humanised animals or transgenic animals to model human diseases, in somatic or germline gene therapy, to generate a transcript image of a tissue or cell type, for detecting differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier or affected individuals, and as hybridisation probes for mapping naturally occurring genomic sequences. PKIN is useful

CC	in southern or northern analysis, dot blot or other membrane-based technologies, in PCR technologies, in dipstick, pin, multiformat enzyme linked immunosorbent (ELISA)-like assays and in microarrays utilising fluids or tissues from patients to detect altered PKIN expression. The present sequence is human PKIN-5
XX	
XX	
QQ	Sequence 645 AA;
SS	Query Match
TT	Best Local Similarity 96.3%; Score 2408; DB 5; Length 645;
UU	Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
VV	
WW	
YY	
ZZ	
AA	
BB	
CC	
DD	
EE	
FF	
GG	
HH	
II	
JJ	
KK	
LL	
MM	
NN	
OO	
PP	
QQ	
RR	
SS	
TT	
UU	
VV	
WW	
XX	
YY	
ZZ	
AAA	
BBB	
CCC	
DDD	
EEE	
FFF	
GGG	
HHH	
III	
JJJ	
KKK	
LLL	
MMM	
NNN	
OOO	
PPP	
QQQ	
RRR	
SSS	
TTT	
UUU	
VVV	
WWW	
XXX	
YYY	
ZZZ	
AAAA	
BBBB	
CCCC	
DDDD	
EEEE	
FFFF	
GGGG	
HHHH	
IIII	
JJJJ	
KKKK	
LLLL	
MMMM	
NNNN	
OOOO	
PPPP	
QQQQ	
RRRR	
SSSS	
TTTT	
UUUU	
VVVV	
WWWW	
XXXX	
YYYY	
ZZZZ	
AAAAA	
BBBBB	
CCCCC	
DDDDD	
EEEEE	
FFFFFF	
GGGGG	
HHHHH	
IIIII	
JJJJJ	
KKKKK	
LLLLL	
MMMMM	
NNNNN	
Ooooo	
PPPPP	
QQQQQ	
RRRRR	
SSSSS	
TTTTT	
UUUUU	
VVVVV	
WWWWW	
XXXXX	
YYYYY	
ZZZZZ	
AAAAA	
BBBBB	
CCCCC	
DDDDD	
EEEEE	
FFFFFF	
GGGGG	
HHHHH	
IIIII	
JJJJJ	
KKKKK	
LLLLL	
MMMMM	
NNNNN	
Ooooo	
PPPPP	
QQQQQ	
RRRRR	
SSSSS	
TTTTT	
UUUUU	
VVVVV	
WWWWW	
XXXXX	
YYYYY	
ZZZZZ	
AAAAA	
BBBBB	
CCCCC	
DDDDD	
EEEEE	
FFFFFF	
GGGGG	
HHHHH	
IIIII	
JJJJJ	
KKKKK	
LLLLL	
MMMMM	
NNNNN	
Ooooo	
PPPPP	
QQQQQ	
RRRRR	
SSSSS	
TTTTT	
UUUUU	
VVVVV	
WWWWW	
XXXXX	
YYYYY	
ZZZZZ	
AAAAA	
BBBBB	
CCCCC	
DDDDD	
EEEEE	
FFFFFF	
GGGGG	
HHHHH	
IIIII	
JJJJJ	
KKKKK	
LLLLL	
MMMMM	
NNNNN	
Ooooo	
PPPPP	
QQQQQ	
RRRRR	
SSSSS	
TTTTT	
UUUUU	
VVVVV	
WWWWW	
XXXXX	
YYYYY	
ZZZZZ	
AAAAA	
BBBBB	
CCCCC	
DDDDD	
EEEEE	
FFFFFF	
GGGGG	
HHHHH	
IIIII	
JJJJJ	
KKKKK	
LLLLL	
MMMMM	
NNNNN	
Ooooo	
PPPPP	
QQQQQ	
RRRRR	
SSSSS	
TTTTT	
UUUUU	
VVVVV	
WWWWW	
XXXXX	
YYYYY	
ZZZZZ	
AAAAA	
BBBBB	
CCCCC	
DDDDD	
EEEEE	
FFFFFF	
GGGGG	
HHHHH	
IIIII	
JJJJJ	
KKKKK	
LLLLL	
MMMMM	
NNNNN	
Ooooo	
PPPPP	
QQQQQ	
RRRRR	
SSSSS	
TTTTT	
UUUUU	
VVVVV	
WWWWW	
XXXXX	
YYYYY	
ZZZZZ	
AAAAA	
BBBBB	
CCCCC	
DDDDD	
EEEEE	

DR N-PSDB; ADC51185.  
 XX Novel DNA or RNA coding a cell-cycle related protein which has nuclear  
 PT export function, useful for screening substance that prevent or treat  
 PT cell cycle abnormality diseases e.g. immunological disease.  
 XX  
 XX  
 PS Disclosure; SEQ ID NO 12; 41pp; Japanese.  
 XX  
 XX The present invention relates to novel cell-cycle related protein such as  
 CC NIMA (Never-in Mitosis, gene A)-related protein kinase of Nek 9 and  
 CC coding sequences such as a cell-cycle related protein (ADC51176) having  
 CC nuclear export function; cell-cycle related protein (ADC51178) having  
 CC nuclear-cytoplasm transport function; cell-cycle related protein  
 CC (ADC51180) having nuclear export function, cytotoxic function and  
 CC transfer function in the nucleus; and/or cell-cycle related protein  
 CC (ADC51182) having transfer function and cytotoxic function in nucleus,  
 CC nuclear export function and nuclear cytoplasm transport function. The  
 CC coding sequences for these proteins are given in ADC51175, ADC51177,  
 CC ADC51179 and ADC51181. The sequences of the invention are useful for  
 CC screening a substance which promotes or suppresses the transfer function  
 CC or cytotoxic function in the nucleus, nuclear-cytoplasm transport  
 CC function and/or a cell-cycle control function. The sequences are also  
 CC useful for treating or diagnosing an immunological disease, neurological  
 CC disease or cancer. The present sequence was used to illustrate the  
 CC invention.  
 XX  
 SQ Sequence 645 AA;

Query Match 96.3%; Score 2408; DB 7; Length 645;  
 Best Local Similarity 100.0%; Pred. No. 6.2e-177; Indels 0; Gaps 0;  
 Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MLKFEAAKCVSGSTAISTYPTKTLIARYVLOQKLGSGFGTVYVSDKAKRGEELKVL 60  
 Db 1 MLKFEAAKCVSGSTAISTYPTKTLIARYVLOQKLGSGFGTVYVSDKAKRGEELKVL 60  
 QY 61 KEISVGLNPNETVQANLEAQLSKLDHPAIVKPHASFEQDNFCIITEYCEGRDLDDKI 120  
 Db 61 KEISVGLNPNETVQANLEAQLSKLDHPAIVKPHASFEQDNFCIITEYCEGRDLDDKI 120  
 QY 121 QEYKQAGKIPFENQIIEWFTIQLLGVDMYHERRILHRDLKSNVFLKNNLLKIGDFGVSR 180  
 Db 121 QEYKQAGKIPFENQIIEWFTIQLLGVDMYHERRILHRDLKSNVFLKNNLLKIGDFGVSR 180  
 QY 181 LMGSDDLATTLTGTHYNSPEALKHQGYDTKSDIWSLACILYEMCCMHAFAGSNFLSI 240  
 Db 181 LMGSDDLATTLTGTHYNSPEALKHQGYDTKSDIWSLACILYEMCCMHAFAGSNFLSI 240  
 QY 241 VLKIVSGDTPSLPERYPKELNAMESMLKNPFLRPSAIEILKIPLYDEQLNLCRYSE 300  
 Db 241 VLKIVSGDTPSLPERYPKELNAMESMLKNPFLRPSAIEILKIPLYDEQLNLCRYSE 300  
 QY 301 MTLEDKXLDCKQEAHIINAMQKRIHLQTLRALSEVQKMTPRERMLRKLQAADKAKKL 360  
 Db 301 MTLEDKXLDCKQEAHIINAMQKRIHLQTLRALSEVQKMTPRERMLRKLQAADKAKKL 360  
 QY 361 KKTIVSKYENSGMQLSRNFQQLSVDLVLEKTHLKGMEKEKEQPEGRILSCSPQDEDE 420  
 Db 361 KKTIVSKYENSGMQLSRNFQQLSVDLVLEKTHLKGMEKEKEQPEGRILSCSPQDEDE 420  
 QY 421 ERWQGRREESDEPTLENLPESQIPSMDLHELESIVEDATSDLGX 466  
 Db 421 ERWQGRREESDEPTLENLPESQIPSMDLHELESIVEDATSDLGX 466

RESULT 5  
 ABG31081  
 ID ABG31081 standard; protein; 645 AA.  
 XX  
 XX ABG31081;  
 XX 21-OCT-2002 (first entry)  
 DT  
 XX

DE Human novel serine/threonine serine kinase.  
 XX  
 KW Human; enzyme; serine/threonine protein kinase; inflammation; cancer;  
 KW arteriosclerosis; psoriasis; SNF kinase; immunogen; chromosome 3.  
 OS  
 XX Homo sapiens.  
 XX  
 PN US2002082189-A1.  
 XX  
 PD 27-JUN-2002.  
 XX  
 PF 07-DEC-2000; 2000US-00731231.  
 XX  
 PR 07-DEC-2000; 2000US-00731231.  
 XX  
 PA (GUEG/) GUEGLER K.  
 PA (KSTC/) KETCHUM K A.  
 PA (DPERA/) DI FRANCESCO V.  
 PA (BEAS/) BEASLEY E M.  
 XX  
 PI Guegler K, Ketchum KA, Di Francesco V, Beasley EM;  
 XX WPI; 2002-598989/54.  
 DR N-PSDB; ABK89295, ABK89296.  
 DR  
 XX New isolated human kinase peptide for detecting a modulator of the  
 PT peptide's expression, activity or function, that can be used to treat  
 PT disorders or disease.  
 XX  
 PS Claim 1; Fig 2; 321pp; English.  
 CC The invention relates to an isolated kinase peptide, comprising, a novel  
 CC human serine/threonine protein kinase sequence, its allelic variant or  
 CC orthologue, where the variant is encoded by a nucleic acid molecule that  
 CC hybridises under stringent conditions to the opposite strand of kinase  
 CC cDNA or gene, or a fragment comprising 10 contiguous amino acids. Also  
 CC included are an antibody that selectively binds the kinase, a gene chip  
 CC comprising a nucleic acid (or its complement) which encodes the kinase, a  
 CC transgenic non-human animal comprising the nucleic acid, a nucleic acid  
 CC vector comprising the nucleic acid and a host cell comprising the vector.  
 CC The kinases used to identify a modulator of the kinase or to identify  
 CC an agent that binds to the kinase, which can be used to treat a disease  
 CC or condition e.g. inflammation, cancer, arteriosclerosis and psoriasis.  
 CC The nucleic acid encoding the kinase is used to produce the kinase. A  
 CC detection agent is used to detect the kinase and an oligonucleotide is  
 CC used to detect a nucleic acid encoding the kinase in a sample. The kinase  
 CC can be used: (a) to raise antibodies against the kinase or to elicit  
 CC another immune response; (b) as a reagent (including a labeling reagent)  
 CC in assays to determine levels of a kinase (or its binding partner or  
 CC ligand) in biological fluids; and (c) as markers for tissues in which the  
 CC corresponding kinase is expressed. The kinase and antibodies against it  
 CC can be used in pharmacogenomic analysis. The kinase can be used to treat  
 CC a disorder characterised by an absence of, inappropriate, or unwanted  
 CC expression of the kinase. The kinase, homologous to SNF kinases (not  
 CC defined) is expressed by a gene located on human chromosome 3. The  
 CC present sequence represents the novel human kinase  
 XX  
 SQ Sequence 645 AA;

Query Match 96.2%; Score 2405; DB 5; Length 645;  
 Best Local Similarity 99.8%; Pred. No. 1.1e-176;  
 Matches 465; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MLKFEAAKCVSGSTAISTYPTKTLIARYVLOQKLGSGFGTVYVSDKAKRGEELKVL 60  
 Db 1 MLKFEAAKCVSGSTAISTYPTKTLIARYVLOQKLGSGFGTVYVSDKAKRGEELKVL 60  
 QY 61 KEISVGLNPNETVQANLEAQLSKLDHPAIVKPHASFEQDNFCIITEYCEGRDLDDKI 120  
 Db 61 KEISVGLNPNETVQANLEAQLSKLDHPAIVKPHASFEQDNFCIITEYCEGRDLDDKI 120  
 QY 121 QEYKQAGKIPFENQIIEWFTIQLLGVDMYHERRILHRDLKSNVFLKNNLLKIGDFGVSR 180



Db 121 QEVQAGKIPFNQIIEWFIQLLLGVYDYMERRILHRDLKSKNVFLKNNLLKIGDFGVS 180  
Qy 181 LLMGSCDLATLTGTPTPHYSPEALKHGGYDTKSDIWSLACILYEMCMNHAFAGSNFLSI 240  
Db 181 LLMGSCDLATLTGTPTPHYSPEALKHGGYDTKSDIWSLACILYEMCMNHAFAGSNFLSI 240  
Qy 241 VLKIVEGDTPSLPERYPKELNAMESMLNKNPSLRPSAIEILKIPYLDEQLNLMCRYSE 300  
Db 241 VLKIVEGDTPSLPERYPKELNAMESMLNKNPSLRPSAIEILKIPYLDEQLNLMCRYSE 300  
Qy 301 MTLEDKNLDCQKEAAHIINAMQKRIHLQTLRALSEVQKTPRERMRLKLAQADEKARKL 360  
Db 301 MTLEDKNLDCQKEAAHIINAMQKRIHLQTLRALSEVQKTPRERMRLKLAQADEKARKL 360  
Qy 361 KKIIVEKYENSKEMOELSRNFQQLSDVVLHEKTHLKGMEKEQPEGRSLSCSPQDEDE 420  
Db 361 KKIIVEKYENSKEMOELSRNFQQLSDVVLHEKTHLKGMEKEQPEGRSLSCSPQDEDE 420  
Qy 421 ERWQGREESDEPTLENLPESQIPSMDLHELESIVEDATSDLGYN 466  
Db 421 ERWQGREESDEPTLENLPESQIPSMDLHELESIVEDATSDLGYN 466

## RESULT 6

ADC99091  
ID ADC99091 standard; protein; 609 AA.  
XX  
AC ADC99091;  
XX  
DT 01-JAN-2004 (first entry)  
XX  
DE Human KPP protein - SEQ ID 44.  
XX  
Kw anti-HIV; antiallergic; antiinflammatory; antianaemic; antiparkinsonian;  
Kw neutropic; anticonvulsant; antiarteriosclerotic; antisthmatic;  
Kw immunosuppressive; antithyroid; cytostatic; hepatotropic; dermatological;  
Kw osteopathic; nephrotropic; angiot; thyromimetic; neuroprotective;  
Kw uropathic; antihelminthic; antiparasitic; antihelminthic; antiparasitic;  
Kw virucide; protozoacide; fungicide; kinase; phosphatase; KPP;  
Kw cell proliferative disorder; atherosclerosis; cirrhosis; hepatitis;  
Kw cancer; developmental; mental retardation; neurological;  
Kw Alzheimer's disease; Parkinson's; autoimmune; inflammatory; Crohn's;  
Kw diabetes mellitus; viral; bacterial; fungal; parasitic; protozoan;  
Kw helminthic infection; transgenic; gene therapy; human; enzyme.  
XX  
Os Homo sapiens.  
XX  
XN WO2003033680-A2.  
XX  
PD 24-APR-2003.  
XX  
PF 17-OCT-2002; 2002WO-US033723.  
XX  
PR 19-OCT-2001; 2001US-0345474P.  
PR 02-NOV-2001; 2001US-0343910P.  
PR 13-NOV-2001; 2001US-0333098P.  
PR 16-NOV-2001; 2001US-0332424P.  
PR 30-NOV-2001; 2001US-0334289P.  
XX  
PA (INCY-) INCYTE GENOMICS INC.  
XX  
PI Bandnan O, Baughn MR, Borowsky ML, Duggan BM;  
PI Emerling BM, Forsythe IJ, Gandhi AR, Gorvad AE, Griffin JA;  
PI Gururajan R, Hafalia AJA, Khan PA, Lal PG, Lee EA, Lee SY;  
PI Lindquist EA, Lu DAM, Lu Y, Marquis JP, Nguyen DB, Arvizu CS;  
PI Ramkumar J, Recipon SA, Richardson TW, Swarnakar A, Tang YT;  
PI Thornton MB, Tran UK, Chawla NK, Warren BA, Yang J, Yao MG, Yue H;  
PI Zebarian Y;  
XX  
DR WPI; 2003-403214/38.  
DR N-PSDB; ADC99143.  
XX

PT New human kinases and phosphatases and polynucleotides, useful for  
PT diagnosing, treating or preventing autoimmune or inflammatory disorders  
PT (e.g. AIDS, allergy or anemia), multiple sclerosis, osteoarthritis,  
XX cancer or hepatitis.  
PS Claim 1; SEQ ID NO 44; 424pp; English.  
XX  
CC The invention relates to a novel isolated polypeptide which is a human  
CC kinase and phosphatase (KPP). The KPP polypeptides, polynucleotides,  
CC agonists and antagonists are useful for diagnosing, treating or  
CC preventing cell proliferative disorders such as atherosclerosis,  
CC cirrhosis, hepatitis and cancer, developmental disorders e.g. mental  
CC retardation, neurological disorders including Alzheimer's disease and  
CC Parkinson's disease, autoimmune and inflammatory disorders such as  
CC Crohn's disease and diabetes mellitus and finally, viral, bacterial,  
CC fungal, parasitic, protozoan or helminthic infections. Furthermore, the  
CC polynucleotides encoding KPP may be useful for creating transgenic  
CC animals to model human disease, as well as during gene therapy  
CC procedures. The current sequence is that of the human KPP protein of the  
CC invention.  
XX  
SQ Sequence 609 AA;

Query Match 87.2%; Score 2182; DB 7; Length 609;  
Best Local Similarity 92.3%; Pred. No. 1.5e-159;  
Matches 430; Conservative 0; Mismatches 0; Indels 36; Gaps 1;  
Qy 1 MLKFEAAKCVSGSTAISTYPTKTLIARRYVLQQLGSGSFGTVYLVSDKAKRGEELKVL 60  
Db 1 MLKFEAAKCVSGSTAISTYPTKTLIARRYVLQQLGSGSFGTVYLVSDKAKRGEELKVL 60  
Qy 61 KEISVGEINPNETVOANLEAQLSKLDHPAIVKPHASFEQDNFCIITECEGRDLDDKI 120  
Db 61 KEISVGEINPNETVOANLEAQLSKLDHPAIVKPHASFEQDNFCIITECEGRDLDDKI 120  
Qy 121 QEVQAGKIPFNQIIEWFIQLLLGVYDYMERRILHRDLKSKNVFLKNNLLKIGDFGVS 180  
Db 121 QEVQAGKIPFNQIIEWFIQLLLGVYDYMERRILHRDLKSKNVFLKNNLLKIGDFGVS 180  
Qy 181 LLMGSCDLATLTGTPTPHYSPEALKHGGYDTKSDIWSLACILYEMCMNHAFAGSNFLSI 240  
Db 181 LLMGSCDLATLTGTPTPHYSPEALKHGGYDTKSDIWSLACILYEMCMNHAFAGSNFLSI 240  
Qy 241 VLKIVEGDTPSLPERYPKELNAMESMLNKNPSLRPSAIEILKIPYLDEQLNLMCRYSE 300  
Db 241 VLKIVEGDTPSLPERYPKELNAMESMLNKNPSLRPSAIEILKIPYLDEQLNLMCRYSE 300  
Qy 301 MTLEDKNLDCQKEAAHIINAMQKRIHLQTLRALSEVQKTPRERMRLKLAQADEKARKL 360  
Db 301 MTLEDKNLDCQKEAAHIINAMQKRIHLQTLRALSEVQKTPRERMRLKLAQADEKARKL 360  
Qy 361 KKIIVEKYENSKEMOELSRNFQQLSDVVLHEKTHLKGMEKEQPEGRSLSCSPQDEDE 420  
Db 361 KKIIVEKYENSKEMOELSRNFQQLSDVVLHEKTHLKGMEKEQPEGRSLSCSPQDEDE 420  
Qy 421 ERWQGREESDEPTLENLPESQIPSMDLHELESIVEDATSDLGYN 466  
Db 393 -----ESDEPTLENLPESQIPSMDLHELESIVEDATSDLGYN 430

## RESULT 7

ADC99090  
ID ADC99090 standard; protein; 540 AA.  
XX  
AC ADC99090;  
XX  
DT 01-JAN-2004 (first entry)  
XX  
DE Human KPP protein - SEQ ID 43.  
XX  
Kw anti-HIV; antiallergic; antiinflammatory; antianaemic; antiparkinsonian;  
Kw neutropic; anticonvulsant; antiarteriosclerotic; antisthmatic;  
Kw immunosuppressive; antithyroid; cytostatic; hepatotropic; dermatological;



CC identifying expressed genes or for physical mapping of human genome. The  
CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight  
CC markers, as a food supplement, for generating antibodies, in medical  
CC imaging, screening and diagnostic assays and for treating cell  
CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's  
CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,  
CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,  
CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver  
CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),  
CC arthritis, etc. Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ

Sequence 497 AA;

Query Match 66.1%; Score 1652; DB 5; Length 497;  
Best Local Similarity 100.0%; Pred. No. 8.8e-119; Indels 0; Gaps 0;  
Matches 318; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 149 MHERRILHRLDKSNVFLKNNLLKIGDFGVSRLLMGSCDLATTLTGTHYMSPEALKEQG 208  
DB 1 MHERRILHRLDKSNVFLKNNLLKIGDFGVSRLLMGSCDLATTLTGTHYMSPEALKEQG 60  
QY 209 YTKSDIWSLACILYEMCCMNAHAFAGSNFLSIVLXIVGDTSPSLPERYKELNATMESNL 268  
DB 61 YTKSDIWSLACILYEMCCMNAHAFAGSNFLSIVLXIVGDTSPSLPERYKELNATMESNL 120  
QY 269 NKNPSLRPSAIFILKIPYLDEQLNLMCRYSEMTLEDKNLDCQKEAAHIINAMQKRIHLQ 328  
DB 121 NKNPSLRPSAIFILKIPYLDEQLNLMCRYSEMTLEDKNLDCQKEAAHIINAMQKRIHLQ 180  
QY 329 TLRALSEVQKMTPRERMRLKLOADEKARKLKKIVVEKYENSXRMQELSRNFQQLSV 388  
DB 181 TLRALSEVQKMTPRERMRLKLOADEKARKLKKIVVEKYENSXRMQELSRNFQQLSV 240  
QY 389 DVLHEKTHLKGMEKEEPEGRSLSCSPQDEDEERWQGREESDEPTLENLPESQIPSPMD 448  
DB 241 DVLHEKTHLKGMEKEEPEGRSLSCSPQDEDEERWQGREESDEPTLENLPESQIPSPMD 300  
QY 449 LHELESIVEDATSDLGYN 466  
DB 301 LHELESIVEDATSDLGYN 318

RESULT 9

ABB72299  
ID ABB72299 standard; protein; 467 AA.  
AC ABB72299;  
DT 04-APR-2002 (first entry)  
DE Rat protein isolated from skin cells SEQ ID NO: 513.

XX Human; rat; mouse; skin cell; skin wound; cancer; growth defect;  
KW developmental defect; inflammatory disease; dermatological; vulnary;  
KW immunomodulator; anti-inflammatory; cytostatic; neuroprotective.  
OS Rattus sp.

XX WO200190357-A1.

PN 29-NOV-2001.

XX 24-MAY-2001; 2001WO-NZ000099.

PF 24-MAY-2000; 2000US-0206650P.

PR 25-JUL-2000; 2000US-0221232P.

XX (GENE-) GENESIS RES & DEV CORP LTD.

PA Watson JD, Strachan L, Sleeman M, Onrust R, Murison JG;  
PI Kumble KD;

XX WPI; 2002-122020/16.  
DR New polynucleotides and polypeptides encoded by the polynucleotides  
XX isolated from skin cells, useful for treating skin wounds, cancers,  
PT growth and developmental defects, inflammatory diseases, or for  
PT modulating immune responses.  
XX  
PS Claim 4; Page 318-319; 466pp; English.

XX The present invention provides the protein and coding sequences of cDNAs  
CC isolated from human, murine and rat skin cell libraries. The sequences  
CC can be used in the development of therapeutic agents useful in the  
CC treatment of skin diseases, including skin wounds, cancer, growth  
CC defects, developmental defects and inflammatory diseases. The proteins  
CC have important roles in the induction of hair growth, cell proliferation  
CC and cell-cell interaction, in maintaining tissue integrity, in wound  
CC healing and in modulating immune responses. The present sequence is a  
CC polypeptide of the invention  
XX

Sequence 467 AA;

Query Match 59.6%; Score 1491; DB 5; Length 467;  
Best Local Similarity 64.3%; Pred. No. 2.1e-106; Indels 80; Gaps 5;  
Matches 301; Conservative 40; Mismatches 47;  
QY 1 MLKFQEAACVSGSSTAISTYPTKILARIYVLOOKLGGSGFTVLYVSDKAKRGEELKVL 60  
DB 1 MLKFQEAACVSGSSTAISTYPTKILARIYVLOOKLGGSGFTVLYVSDKAKRGEELKVL 58  
QY 61 KEISVGEINNETVQANLEAQLSKLDHPAIVKHFASFEQDNFCIITEYCEGRDLDKI 120  
DB 59 KEISVGEINNETVQANLEAQLSKLDHPAIVKHFASFEQDNFCIITEYCEGRDLDKI 118  
QY 121 QEYKQAGKIPFNCIIEWFQILLGVVDYMERHRLDKSNVFLKNNLLKIGDFGVS 180  
DB 119 QEYKQAGKIPFNCIIEWFQILLGVVDYMERHRLDKSNVFLKNNLLKIGDFGVS 178  
QY 181 LMGSCDLATTLTGTHYMSPEALKEQGDTKSDIWSLACILYEMCCMNAHAFAGSNFLSI 240  
DB 179 LMGSCDLATTLTGTHYMSPEALKEQGDTKSDIWSLACILYEMCCMNAHAFAGSNFLSI 238  
QY 241 VLKIVEGDTSPSLPERYKELNATMESMLNKNPSLRPSAIFILKIPYLDEQLNLMCRYSE 300  
DB 239 VLKIVEGDTSPSLPERYKELNATMESMLNKNPSLRPSAIFILKIPYLDEQLNLMCRYSE 298  
QY 301 MTLEDK-NLDCQKEAAHIINAMQKRIHLQTLRALSEVQKMTPRERMRLKLOADEKARK 359  
DB 299 ATLEDKNSACQKEAAHVNAI-----  
QY 360 LKXIVE-EKYENSXRMQELSRNFQQLSVLXIVGDTSPSLPERYKELNATMESMLNKNPSLRPSAIFILKIPYLDEQLNLMCRYSE 418  
DB 321 --KITEGRRYKNNKRAKELRSNQFSGVSAHVL-----  
QY 419 DEERWQGREESDEPTLENLPESQIPSPMDLHELESIVEDATSDLGYN 466  
DB 352 -----QELDELITLESLSQPSLPCNLNDELPSLEGITVDLGYN 390

RESULT 10

ADCS1182  
ID ADCS1182 standard; protein; 357 AA.  
XX ADCS1182;  
AC ADCS1182;  
DT 18-DEC-2003 (first entry)  
XX Human cell-cycle related protein, SEQ ID 8.

XX Human; cytostatic; cell-cycle related protein; nuclear export;  
KW nuclear-cytoplasm transport; cytotoxic; cell-cycle control;  
KW immunological disease; neurological disease; cancer.



Db 181 KPNYKSDVWALGCCVYEMATLKHAFAKMDNSLVTRIIEGKLPAMPDRYSPELAELIRT 240  
 QY 267 MLNKNPSLRPSAIEILKIPYLDQQLNLCRYSEMTELD--KNLDCQ-KEAAHIIN---- 319  
 Db 241 MLSKRPPEPVSILRQPIKQISFFLEATKIKTSKNINKNKNGDSQSFPFATVVSGEAE 300  
 QY 320 AMQKRIHLQTLRA-----LSEVQKTPPRERMLRKL--QAADEKARKLKKIVEEYEN 371  
 Db 301 SNHEVHPQLPPLSEGSQTYIMGEGKCLSQEKPRASGLLSPASLKAHTCKQDL-----SN 355  
 QY 372 SKRMQELSRNFOQLSDVHLHETKHLKGMEE-----KEEQPE-----GRLSCSPQD 417  
 Db 356 TTELATISSVN-----IDIL-----PAKGRDSVSDGFVQENQPRYLDASNELGGI-CSISQ 405  
 QY 418 EDEERWQGREESDEPTLENLPESQIPSMDL 449  
 Db 406 VEEMQLQNTKSAQP--ENL---IPMWSSDI 432

RESULT 12

ADB37570

ID ADB37570 standard; protein; 841 AA.

XX ADB37570;

XX 04-DEC-2003 (first entry)

XX Neural thread protein-related protein #33.

XX Cytostatic; Antitumour; Antipsoriatic; Dermatological;

XX Antiatherosclerotic; Antiarteriosclerotic; Vasotropic; Antiinflammatory;

XX Immunosuppressive; Tranquillizer; Antileptic; Virucide; AD7C-NTP;

XX neural thread protein; neuritic sprouting.

XX Unidentified.

XX WO2003008444-A2.

XX 30-JAN-2003.

XX 19-JUL-2002; 2002WO-CA001106.

XX 19-JUL-2001; 2001US-0306150P.

XX 19-JUL-2001; 2001US-0306161P.

XX 16-NOV-2001; 2001US-0331477P.

XX (NYMO-) NYMOX CORP.

XX Averbach PA, Gemmell J;

XX WPI; 2003-248000/24.

XX Novel Related peptide or AD7C-neural thread peptide, useful for treating

XX unwanted cellular proliferations, glandular hyperplasia, unwanted facial

XX hair, warts and unwanted fatty tissue.

XX Claim 1; Page 34; 109pp; English.

XX The present invention relates to AD7C-neural thread protein (NTP) and

XX related proteins and peptides (1; ADB37528-ADB37641). The sequences are

XX useful for treating a condition in a patient requiring removal or

XX destruction of cells. The condition can be selected from benign or

XX malignant tumour of a tissue, hyperplasia, hypertrophy or overgrowth of a

XX tissue, virally, bacterially or parasitically altered tissue, or

XX malformation of a tissue, where the tissue is selected from lung, breast,

XX stomach, pancreas, prostate, bladder, bone, ovary, skin, kidney, sinus,

XX colon, intestine, stomach, rectum, oesophagus, heart, spleen, salivary

XX gland, blood, brain and its coverings, spinal cord and its coverings,

XX muscle, connective tissue, adrenal, parathyroid, thyroid, uterus, testis,

XX pituitary, reproductive organs, liver, gall bladder, eye, ear, nose,

XX throat, tonsils, mouth, lymph nodes and lymphoid tissue. The condition is

XX preferably tonsillar hypertrophy, prostatic hyperplasia, psoriasis,

XX eczema, dermatosis, cosmetic modification to a tissue (skin, eye, ear,

XX

CC nose, throat, mouth, muscle, connective, hair or breast tissue), vascular  
 CC disease (atherosclerosis or arteriosclerosis), haemorrhoids, varicose  
 CC veins, inflammatory disease, autoimmune disease, metabolic disease,  
 CC hereditary/genetic disease, traumatic disease or physical injury,  
 CC nutritional deficiency disease, infectious disease, amyloid disease,  
 CC fibrosis disease, storage disease, congenital malformation, enzyme  
 CC deficiency disease, poisoning, intoxication, environmental disease,  
 CC radiation disease, endocrine disease, degenerative disease and mechanical  
 CC disease. The peptides are useful for treating unwanted cellular  
 CC proliferations, glandular (e.g. prostate) hyperplasia, unwanted facial  
 CC hair, warts and unwanted fatty tissue, or for preparing antibodies that  
 CC recognize and/or bind to Related proteins, Related peptides or NTP  
 CC peptides.

XX Sequence 841 AA;

XX Query Match 24.5%; Score 613; DB 7; Length 841;

XX Best Local Similarity 34.7%; Pred. No. 2.3e-38;

XX Matches 157; Conservative 86; Mismatches 133; Indels 56; Gaps 17;

QY 29 YVLOQKLGSGFTYVYLVSDKKAKRGBELKVLKBEISVGNELNPNETVOANLEAQLSKLDH 88

Db 6 YCYLRVVGKSGYGEVTLV--KHRRDGKQYVVKLNLNRNASSRRRAAEQAQLLSQLKH 62

QY 89 PAIVKEHASFVEQDNFC-IITEYCEGEDLDDKICEYKQAGKIPENQIIEFWIQLLLGVD 147

Db 63 PNIVTYKESWEGDGLLYIVNGCEGDLRYLKE--QKGLLPENOVVFWFOIAVALQ 120

QY 148 YMERRILHRDLKSKNVFL-KNNLLKIGFVGRLLMGSCDLATTLTGTTHYMSPEALKH 206

Db 121 YLHEKHILHRDLKTONVFLRTNIIKVGDLGIARVLENHCDMASTLIGTFYVNSPELFSN 180

QY 207 QGYDTKSDIWSLACILYEMCCMNAFAGSNFLSVLKIVGEDTPSLPERYPKELNAMES 266

Db 181 KPNYKSDVWALGCCVYEMATLKHAFAKMDNSLVTRIIEGKLPPEPRDYSPELAELIRT 240

QY 267 MLNKNPSLRPSAIEILKIPYLDQQLNLCRYSEMTELD--KNLDCQ-KEAAHIIN---- 319

Db 241 MLSKRPPEPVSILRQPIKQISFFLEATKIKTSKNINKNKNGDSQSFPFATVVSGEAE 300

QY 320 AMQKRIHLQTLRA-----LSEVQKTPPRERMLRKL--QAADEKARKLKKIVEEYEN 371

Db 301 SNHEVHPQLPPLSEGSQTYIMGEGKCLSQEKPRASGLLSPASLKAHTCKQDL-----SN 355

QY 372 SKRMQELSRNFOQLSDVHLHETKHLKGMEE-----KEEQPE-----GRLSCSPQD 417

Db 356 TTELATISSVN-----IDIL-----PAKGRDSVSDGFVQENQPRYLDASNELGGI-CSISQ 405

QY 418 EDEERWQGREESDEPTLENLPESQIPSMDL 449

Db 406 VEEMQLQNTKSAQP--ENL---IPMWSSDI 432

RESULT 13

ABUS4632

ID ABUS4632 standard; protein; 489 AA.

XX ABUS4632;

XX 03-JUN-2003 (first entry)

XX Human NOVX polypeptide #91.

XX Human; NOVX; metabolic disorder; cardiomyopathy; diabetes; ASD;

XX hypertension; congenital heart defect; aortic stenosis; valve disease;

XX atrial septal defect; atrioventricular canal defect; ductus arteriosus;

XX pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;

XX tubercous sclerosis; scleroderma; atherosclerosis; infectious disease;

XX obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;

XX Parkinson's disease; immune disorder; haematopoietic disorder;

XX haemophilia; hypercoagulation; Crohn's disease; cancer.

XX Homo sapiens.

XX CC disorders, Alzheimer's disease, Parkinson's disease, immune disorders,  
PN CC haematopoietic disorders, haemophilia, hypercoagulation, Crohn's disease  
XX CC and cancer. Sequences ABU54542-ABU54647 represent human NOVX polypeptides  
PD CC of the invention  
XX SQ Sequence 489 AA;  
PF Query Match 23.9%; Score 598.5; DB 6; Length 489;  
PR Best Local Similarity 31.8%; Pred. No. 1.5e-37;  
PR Matches 152; Conservative 89; Mismatches 151; Indels 89; Gaps 15;  
QY 29 YVLQQLGSGFVTYLVSDKKAKRGEELKVKELKESVGLNPNETVQANLEQLKLDH 88  
DB 4 YWVLRMIGEGSGRALLVQHSS--NQMFAMKEIRLPKVTN-TQNSRKEAVLLAKMH 59  
QY 89 PAIVKPHASVEQDNFCIITEYCEGEDLDKIQEYKQAGKIPENQIIEFWIQLLGVY 148  
DB 60 PNIVAFKESFEAECHLYVMEYCDGDLQKIKQ--QKGLFPEDQLLWFTQMLGVNH 117  
QY 149 MHERRILHRLDKSNVFL-KNLLKIGDFGVSRLLMGSCDLATTLTGTTHYMSPEALKH 207  
DB 118 IHKRVLRDIKSNIFLTQNGKVKGLDGFSAELLSPMAFACTYVGTTPYVPPEIWE 177  
QY 208 GYTKSDIWSLACILYEMCMNHAFAGSNFLSIVLIVGSDTSPILPERYPKELNAMES 267  
DB 178 PYNNKSDIWSLGLCYELCTLKHFPQANSWKNLKVCQGC-SPLSHYSYELQFLVKOM 237  
QY 266 LKNPRLPSAIRILKIPYDEOLQ-----NLMCRYSEMTLED-KNL-----DC 310  
DB 238 FKRNPSHRPSATTLISRGIVARLVQKLPPEIIMEYGEVLEIKNSKENTPRKKQEE 297  
QY 311 QKEAAHI-INAMQRIHLQTLRALSEVQVQTPRRMRLKQADKAKLKVIEE--- 366  
DB 298 DRKGSHTDLESINENLVESALRRVNEEK--GNKSVHLRKASSPNLHRQWKNVENTAL 355  
QY 367 -----KYEENSKRQEUR-----SRNFQQLS 387  
DB 356 TALENASILTSSLTAEDDRGSVIKSKNTRIQWLKETPDTLLNLKVAJSLAFQTY 415  
QY 388 VDVLHEXTHLKG--MBEKEE-----QPEGLSCSPQDEBERWQGRESEDE 432  
DB 416 IYRPGSGFLKGPLSEBTEASDSDGCHDSVILDPE-RLEPGLDESDTD-----FE 470  
QY 433 P 433  
DB 471 P 471  
RESULT 14  
AAO26613  
ID AAO26613 standard; protein; 774 AA.  
XX AC AAO26613;  
XX DT 28-MAR-2003 (first entry)  
XX DE Serine/threonine protein kinase-like enzyme related mouse protein.  
XX KW Cytostatic; antiparkinsonian; nootropic; neuroprotective; cardiant;  
KW hypotensive; antiarrhythmic; antianalgesic; analgesic; antiinflammatory;  
KW antididiabetic; tranquiliser; antitonic; antidepressant; uropathic;  
KW anorectic; serine/threonine protein kinase; enzyme; cancer;  
KW central nervous system disorder; cardiovascular disorder; diabetes;  
KW chronic obstructive pulmonary disease; obesity; genito-urinary system;  
KW pain; genetic testing; mouse.  
XX OS Mus musculus.  
XX PN WO200299096-A1.  
XX PD 12-DEC-2002.  
XX PF 06-JUN-2002; 2002WO-EP006203.

XX GUO X, Kekuda R, Miller CE, Malyankar UM, Spytek KA;  
PI Pattarajan M, Liu X, Gusev VY, Li L, Vernet CAM, Zerhusen BD;  
PI Gorman L, Shency SG, Pena CE, Smithson G, Burgess CE, Gerlach V;  
PI Padigaru M, Shimkets RA, Gangalli EA, Taupier RJ, Casman SJ, Ji W;  
PI Anderson DW, Leite SW, Rastelli L, Edinger SR, Stone DJ;  
PI Macdougall JR, Rothenberg ME, Mazur A, Millet I, Peyman JA;  
PI Ellerman K;  
XX WPI; 2003-046856/04.  
DR N-PSDB; ABX72260.  
XX New isolated NOVX polypeptide useful for treating atherosclerosis,  
PT metabolic disorders, diabetes, obesity, infectious disease, anorexia,  
PT neurodegenerative disorders, Alzheimer's disease and cancer.  
XX Claim 1; Page 288; 66pp; English.  
XX The invention relates to human polypeptides, termed NOVX, and the  
CC polynucleotides encoding them. The polypeptides and polynucleotides are  
CC useful for diagnosing disease, and screening for potential therapeutic  
CC agents. The sequences are useful for treating metabolic disorders,  
CC cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic  
CC stenosis, atrial septal defect (ASD), atriocentric stenosis, ventricular  
CC ductus arteriosus, pulmonary stenosis, subaortic stenosis, scleroderma,  
CC septal defect (VSD), valve diseases, tuberculous atherosclerosis, atherosclerosis, obesity, infectious disease, anorexia, neurodegenerative

XX 07-JUN-2001; 2001US-0296164P.  
 PR 19-SEP-2001; 2001US-0323100P.  
 PR 25-OCT-2001; 2001US-0330578P.  
 PR 17-JAN-2002; 2002US-0348601P.  
 XX (FARB ) BAYER AG.  
 XX Smolyar A, Horner EJ, Thelwell C;  
 XX WPI; 2003-140620/13.  
 XX New human serine/threonine protein kinase-like enzyme polypeptide and  
 PT polynucleotide, useful for regulating the activity of the protein kinase-  
 XX like enzyme to prevent, treat or ameliorate diabetes, cancer or obesity.  
 XX Disclosure; Fig 3; 157pp; English.  
 XX The invention relates to an isolated polynucleotide comprising: a  
 CC polynucleotide encoding a serine/threonine protein kinase-like enzyme  
 CC polypeptide; a sequence of 840, 990, 1120, 2142, or 1383 base pairs fully  
 CC defined in the specification; a polynucleotide that hybridizes to the  
 CC serine/threonine protein kinase-like enzyme DNA or a fragment thereof; or  
 CC a degenerate analogue of the said polynucleotides. The serine/threonine  
 CC protein kinase-like enzyme polypeptide and polynucleotide are useful in  
 CC preventing, ameliorating, or treating diseases associated with serine/  
 CC threonine protein kinase-like enzyme dysfunction such as cancer, central  
 CC nervous system disorders, cardiovascular disorders, chronic obstructive  
 CC pulmonary disease, diabetes, obesity, or disorders of the genito-urinary  
 CC system. These can also be used to treat pain associated with the above  
 CC disorders. The serine/threonine protein kinase-like enzyme polypeptide is  
 CC also useful in diagnostic assays or in genetic testing. The expression  
 CC vector or the reagent is useful in preparing a medicament for modulating  
 CC the activity of a serine/threonine protein kinase-like enzyme in a  
 CC disease, e.g. cardiovascular disorder, obesity, a disorder of the genito-  
 CC urinary system, a central nervous system disorder, diabetes, cancer, or  
 CC chronic obstructive pulmonary disease. The methods are useful in  
 CC producing and detecting the polynucleotide and polypeptide and in  
 CC screening for agents that modulate the activity of the serine/threonine  
 CC protein kinase-like enzyme polypeptide. This sequence represents a mouse  
 CC protein relating to the the serine/threonine protein kinase-like enzyme  
 CC of the invention  
 XX  
 XX Sequence 774 AA;

Query Match 23.9%; Score 598; DB 6; Length 774;  
 Best Local Similarity 28.5%; Pred. No. 3e-37;  
 Matches 141; Conservative 94; Mismatches 132; Indels 128; Gaps 13;  
 QY 28 RYVLOKLGSGSGFTYLVSDKAKGELKYLKELISVCELNPNTVQANLEAQLLSKLD 87  
 DB 3 KVRILQKIGSGFGKAVLV--KSTEDGRHY-VIKELINRMSDKERQSRREAVLANMK 59  
 QY 88 HPAIVKPHASVFEQDNFCITTEYCEGRDLDDKIOEYKQAGKIFPENQIIEFWIQLLQVD 147  
 DB 60 HENIVQYKESPENGSLVIMDYCSGGDLFKRIN--AQKALFQEDQILDNFVQICLAK 117  
 QY 148 YNHERILHRDLKSNVL-KNNLLKIGDGVSRLLMGSCDLATLTCTHYMSPALKH 206  
 DB 118 VVHDKRLHRDLKSNVILKDTGTVOLGFGIARVLNSTVELARTICITPTLYLSPICEIN 177  
 QY 207 QGYDTSIWSIACILYEMCMNHAFAGSNFLSVILKIVEGDTPLPERYPKELNAIMES 266  
 DB 178 KPYNNKSDIWDALGVLYELCTLKHAFAENKMKVLKIIISGFFPVSHYSDYLSLLSQ 237  
 QY 267 MLNKPSPRLPSAIELKLPYLDQIQNLN-----C----- 296  
 DB 238 LFKRNPRDRPSVNSILEKGFIAKRIEFLSPQLIAEIEFCLTKSKFGPQLPGKRPSAGQ 297  
 QY 297 -----RYSEMTELDKNLDCOKEAAHINAMQRIHLQTLR 331  
 DB 298 GVSSFPQAQITKPAKYGVPLTYKYGDKLLEKKPPKQAHQI----- 344

QY 332 ALSEVQKMTPRMELRLKLAQADEKARKLKKIVKYEENSKRMQELSRNFQQLSVTVL 391  
 DB 345 ---PVKQNSGEERKQSEEA---KGRLEFIEKEK-----KQKQIRFLKAEQMK--- 391  
 QY 392 HEKTHLKGMEEKKEQP-----EGRLS---CSPO----- 416  
 DB 392 QEQFLERINRAREQCHRWNLRAGSGEVKASFFGICGAVSPSCSPRGQYEHVHAIFDQ 451  
 QY 417 -----DEDERWQG 425  
 DB 452 MQLRAEDNEARWKG 466  
 XX RESULT 15  
 XX ABP71711  
 XX ID ABP71711 standard; protein; 774 AA.  
 XX AC ABP71711;  
 XX DT 17-APR-2003 (first entry)  
 XX DE NEK-like serine/threonine protein kinase NEK1\_mouse # SEQ ID 3.  
 XX KW NEK-like serine/threonine protein kinase; cytostatic; cardiant;  
 XX antiinflammatory; nootropic; neuroprotective; cancer; colon cancer;  
 XX cardiovascular disorder; diabetes; COPD; CNS disorder; mouse; rat.  
 XX OS Mus musculus.  
 XX Rattus norvegicus.  
 XX WO2003000903-A2.  
 XX 03-JAN-2003.  
 XX 24-JUN-2002; 2002WO-EP006948.  
 XX 25-JUN-2001; 2001US-0300068P.  
 XX 07-DEC-2001; 2001US-0336704P.  
 XX (FARB ) BAYER AG.  
 XX Xiao Y;  
 XX WPI; 2003-184051/18.  
 XX New polynucleotide encoding a NEK-like serine/threonine kinase  
 PT polypeptide useful for treating diseases associated with kinase  
 PT dysfunction, e.g. cardiovascular disorders, cancer such as colon cancer,  
 PT diabetes and CNS disorders.  
 XX Disclosure; Fig 3; 149pp; English.  
 XX The invention relates to a newly isolated polynucleotide encoding an NEK-  
 CC like serine/threonine protein kinase. The activity of the polynucleotide  
 CC and polypeptide of the invention may be described as cytostatic,  
 CC cardiant, antiinflammatory, nootropic and neuroprotective. The expression  
 CC vector and reagent of the invention are useful for the preparation of a  
 CC medicament for modulating the activity of an NEK-like serine/threonine  
 CC kinase in a disease, such as cancer (e.g. colon cancer), cardiovascular  
 CC disorder, diabetes, COPD or CNS disorder. The polypeptides may also be  
 CC used to identify compounds which may act as activators or inhibitors at  
 CC the enzyme's active site, to raise specific antibodies which can block  
 CC the enzyme and effectively reduce its activity, as a bait protein in a  
 CC two-hybrid or three-hybrid assay to identify other proteins which bind to  
 CC or interact with the human NEK-like serine/threonine kinase polypeptide  
 CC and modulate its activity, and for the immunisation of mammals. The  
 CC current sequence represents an NEK-like serine/threonine protein kinase  
 CC designated NEK1\_mouse  
 XX Sequence 774 AA;  
 XX Query Match 23.9%; Score 598; DB 6; Length 774;  
 XX Best Local Similarity 28.5%; Pred. No. 3e-37;





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.  
OM protein - protein search, using sw model  
Run on: October 1, 2004, 11:16:00 ; Search time 33 Seconds  
(without alignments)  
754.052 Million cell updates/sec  
Title: US-10-803-278-4  
Perfect score: 2501  
Sequence: 1 MLKFEAAKCVSGSTAISTY.....LQYHGDCLNLSLDYWKNEK 482  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 389414 seqs, 51625971 residues  
Total number of hits satisfying chosen parameters: 389414  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries  
Database :  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/ECTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.ppep:\*  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	598	23.9	616	4	US-09-873-404-4
2	584	23.4	640	4	US-09-873-404-2
3	495	19.8	445	4	US-09-167-206-2
4	493	19.7	345	3	US-09-173-581-6
5	493	19.7	345	3	US-09-420-915-6
6	454	18.2	692	4	US-09-992-481-2
7	445	17.8	302	4	US-09-579-6648-9
8	443	17.7	302	3	US-09-221-235-2
9	443	17.7	302	3	US-09-221-928-2
10	443	17.7	302	3	US-09-221-527-2
11	443	17.7	302	3	US-09-221-236-2
12	443	17.7	302	3	US-09-221-416-2
13	443	17.7	302	3	US-09-221-245-2
14	443	17.7	302	3	US-09-163-115-2
15	443	17.7	302	3	US-09-221-528-2
16	443	17.7	302	3	US-09-221-553-2
17	443	17.7	302	3	US-09-221-237-2
18	440.5	17.6	699	4	US-09-457-0408-18
19	429.5	17.2	979	3	US-08-870-529-2
20	429.5	17.2	979	4	US-09-544-794-2
21	404.5	16.2	233	2	US-08-712-709-1
22	404.5	16.2	233	3	US-09-111-444-1
23	404.5	16.2	233	3	US-09-541-228-1
24	395.5	15.8	293	3	US-09-049-671-1
25	395.5	15.8	293	3	US-09-295-068-1
26	382.5	15.3	239	2	US-08-712-709-7
27	382.5	15.3	239	2	US-09-049-671-3

SUMMARIES

Query Match	23.9%	Score 598;	DB 4;	Length 616;
Best Local Similarity	28.5%;	Pred. No. 4.3e-39;		
Matches	141;	Conservative	94;	Mismatches 132; Indels 128; Gaps 13;
Qy	28	RYVLQKLGSGSTGTVYLVSDKAKGEEELKVLKISVGLNPNETVOANLEAQLLSKLD	87	
Db	3	KYVRLQKIGSGSGKAVLV--KSTEDGRHY-VIKEINISMSDKERQESREVAVLANK	59	
Qy	88	HPATVKEFHAFVQCDNFCIITEYCEGRDLDDKIOEVKQAGKIFPENQIIEWFTQLLIGVD	147	
Db	60	HPNIVQYKESFENGSLYIVNDYCEGGDLFKRN--AQKALFQEDQILDWFIQICLAK	117	
Qy	148	YMERRILHRLDKSKNVFL-KNNLLKIGDFGVSRLLMGSCDLATTITGTTHYVSPEALKH	206	
Db	118	HVHDKRLHRLDIKSONIFLTQDGTVOQLGFGIARVLNSTVELARTCTGTPTVLSPEICEN	177	
Qy	207	QGVDTKSDIWSLACILYEMCCMHAFAGSNFLSVLKIIVEGDTPLSPERYPEKELNAMES	266	
Db	178	KPNYKSDIWSLACILYEMCCMHAFAGSNFLSVLKIIVEGDTPLSPERYPEKELNAMES	237	
Qy	267	MLNKNPSLRPSAIEILKIPYLDQLQNLN-----C-----	296	
Db	238	LFKRNPRDRSVNSILEKGFIAKRIEFLSPQIAIEEFLKTLKSGFPQPLPKGRPASGQ	297	
Qy	297	-----RYSENTIEDKNDLDOCKEAAHINAMQKRIHLQTLR	331	
Db	298	GVSSFVPAQKITPAKYGVPLTYKYGDKLLEKPPPKKHQAHOI-----	344	
Qy	332	ALSEVQKMTPRMRRLKLOADEKARKLKKIVEEKVEENSKMQBLRSNFQOLSVDVL	391	

ALIGNMENTS

RESULT 1  
US-09-873-404-4  
; Sequence 4, Application US/05873404  
; Patent No. 650656  
; GENERAL INFORMATION:  
; APPLICANT: WEBSTER, Marion et al  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001212-CIP  
; CURRENT APPLICATION NUMBER: US/09/873,404  
; CURRENT FILING DATE: 2001-06-05  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 616  
; TYPE: PRT  
; ORGANISM: Mouse  
US-09-873-404-4

Query Match	23.9%	Score 598;	DB 4;	Length 616;
Best Local Similarity	28.5%;	Pred. No. 4.3e-39;		
Matches	141;	Conservative	94;	Mismatches 132; Indels 128; Gaps 13;
Qy	28	RYVLQKLGSGSTGTVYLVSDKAKGEEELKVLKISVGLNPNETVOANLEAQLLSKLD	87	
Db	3	KYVRLQKIGSGSGKAVLV--KSTEDGRHY-VIKEINISMSDKERQESREVAVLANK	59	
Qy	88	HPATVKEFHAFVQCDNFCIITEYCEGRDLDDKIOEVKQAGKIFPENQIIEWFTQLLIGVD	147	
Db	60	HPNIVQYKESFENGSLYIVNDYCEGGDLFKRN--AQKALFQEDQILDWFIQICLAK	117	
Qy	148	YMERRILHRLDKSKNVFL-KNNLLKIGDFGVSRLLMGSCDLATTITGTTHYVSPEALKH	206	
Db	118	HVHDKRLHRLDIKSONIFLTQDGTVOQLGFGIARVLNSTVELARTCTGTPTVLSPEICEN	177	
Qy	207	QGVDTKSDIWSLACILYEMCCMHAFAGSNFLSVLKIIVEGDTPLSPERYPEKELNAMES	266	
Db	178	KPNYKSDIWSLACILYEMCCMHAFAGSNFLSVLKIIVEGDTPLSPERYPEKELNAMES	237	
Qy	267	MLNKNPSLRPSAIEILKIPYLDQLQNLN-----C-----	296	
Db	238	LFKRNPRDRSVNSILEKGFIAKRIEFLSPQIAIEEFLKTLKSGFPQPLPKGRPASGQ	297	
Qy	297	-----RYSENTIEDKNDLDOCKEAAHINAMQKRIHLQTLR	331	
Db	298	GVSSFVPAQKITPAKYGVPLTYKYGDKLLEKPPPKKHQAHOI-----	344	
Qy	332	ALSEVQKMTPRMRRLKLOADEKARKLKKIVEEKVEENSKMQBLRSNFQOLSVDVL	391	

```
Db 345 ---PVKKNNSGEEKKVSBEAA--KKRLLEFTEKE-----KQDQIRFLKAEQMK---R 391
QY 392 HEKTHLKGMEKEEQ-----EGRLS---CSPO----- 416
Db 392 QEQRLERINRAEQWRNVLRAAGSGEVKASFFGIGGAVSPSPGQVHYHAIFDQ 451
QY 417 -----DEDEERWQG 425
Db 452 MQLRAEDNEARWKG 466

RESULT 2
US-09-873-404-2
; Sequence 2, Application US/09873404
; Patent No. 6500656
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001212-CIP
; CURRENT APPLICATION NUMBER: US/09/873,404
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Human
US-09-873-404-2

Query Match 23.4%; Score 584; DB 4; Length 640;
Best Local Similarity 29.2%; Pred. No. 5.8e-38;
Matches 149; Conservative 94; Mismatches 164; Indels 104; Gaps 14;

QY 28 RYVLOQKLGSGFGTVLVSDKKAKRGEELKVLKEISVGEINPNETVQANLEAQLSKLD 87
Db 3 KYDVIRAKGAGFGKAYLA---KGRSDSKHCVCVKEINPEKMPIQEAKSKKEVILLERMK 59
QY 88 HPAIVKPHASVFEQDNFCIITEYCEGRDLDKIQEYKQAGKIFPENQIIEWFIQLLQVD 147
Db 60 HPNIVAFNSFGNGRLFTVMEYCGGDLMKRN--RQGVLFSDQILGWVQVQLSLGK 117
QY 148 YHERIRILHRDLKSNVFLKN--LLKIGDFGVSRLLMGSCDLATTLTGTTPHYMSPEALK 205
Db 118 HIHDRILHRDLKSNVFLKN--LLKIGDFGVSRLLMGSCDLATTLTGTTPHYMSPEALK 177
QY 206 HQGYDTKSDIWSLACILYEMCMNHAFAGSNFLSVLKIIVEGDTPSLPERYPKELNAIWE 265
Db 178 NKPNYNTDITWSLGVLYELCTLKHPFEGNNLQQLVLKICQAHFAPISFGFSRELHSLIS 237
QY 266 SMLKNPSPRPSAIEILKIPYLDEQLQNLKCRYSMTLEDKNDLDCQKAAHII----- 318
Db 238 QLFQVSPDRPSINSILKRPF-----LENLIPKY--LTPE-----VQEEFSHMLICEAGAP 287
QY 319 ----NMQKRIHLQTRALSEVQKMTPRERMLR-----KLOADEKARKLK 361
Db 288 ASRHAGKVVQKQIKQVFRFG---KCPPRSRSISVPIKRNAILHRNEWPPAGAKARSIG 344
QY 362 KIVEEK-----YEENSRMQELSRNFQQLSDVVLHEKTHLKG----- 399
Db 345 MIERPKIAAVCGHYDYVYLAQLMLRRRAHKPSVHIPQENTGVEDYQGTTRGPPSQWP 404
QY 400 ---MEKEQEGRLSCSPQ-----DEDEERWQREE-----ESDEPTLENLPESQIPSPMD 448
Db 405 AEYLQKFEAQQVKLVKQGLRPPSASFPNQRQELSRNGEPRFQBLPRK----- 458
QY 449 LHELESIVEDATSDLGHGDCNLISLIDYWK 479
Db 459 -----NEMKQEQYWK 468

RESULT 3
```

```
US-09-167-206-2
; Sequence 2, Application US/09167206A
; Patent No. 6476193
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Schulz, Vincent P.
; APPLICANT: Yang, Mei-Jia
; TITLE OF INVENTION: NIK1 PROTEIN AND NIK1 PROTEIN COMPLEXES
; FILE REFERENCE: 15966-521 NIK1 protein complexes
; CURRENT APPLICATION NUMBER: US/09/167,206A
; CURRENT FILING DATE: 1998-10-06
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-167-206-2

Query Match 19.8%; Score 495; DB 4; Length 445;
Best Local Similarity 32.7%; Pred. No. 3.8e-31;
Matches 139; Conservative 70; Mismatches 168; Indels 48; Gaps 14;

QY 26 ARRYVLOQKLGSGFGTVLVSDKKAKRGEELKVL--KEISVGEINPNETVQANLEAQL 83
Db 5 AEDYEVLYTIGTSGYGR-----QKIRKSDGKILVWKELDYGSMTEAKQMLVSEVNL 59
QY 84 SKLDHPAIVKPHASVFEQDN--FCIITEYCEGRDLDKIQEYKQAGKIFPENQIIEWFIQ 141
Db 60 RELKHPNIVSYDRIIDRTNTTLYIVMEYCEGGDLASVITKTKERQYLDDEFVLKWTQ 119
QY 142 LLGLGVDMYHER-----RIHLRDLKSNVFL--KQNLKIGDFGVSRLLMGSCDLATTLTG 194
Db 120 LTLALKECHRSDDGGHTVLRDLKPNVFLDGQN--VKLGDFGLARILNHDTSFAKTFVG 178
QY 195 TPHYMSPEALKHQGYDTKSDIWSLACILYEMCMNHAFAGSNFLSVLKIIVEGDTPSLPE 254
Db 179 TPYMSPEQNNRNSYNEKSDIWSLGLCLLYELCALMPPTAFSQKELAGKIREGKFRIPY 238
QY 255 RYPKELNAIMESMLNKPSPRPSAIEILKIPYL-----DEQLQNLKCRYSMTLEDKNDL 309
Db 239 RYSDLENEITRMLNLDKDYHRPSVEILENPLIADLVADQRENLERRGQLGEPEKSD 298
QY 310 CQKEAAHINAMQ--KRIHLQ-----TLRALSE-----VQKMTPRERMLRKLQADEKARK 359
Db 299 ----SSPVLSELKLTQLQERERALKARBERLEQKEQELCVREERLAEDKLARAENLLKN 354
QY 360 LKKIVEKYEENSRMQELSRNFQQLSDVVLHEKTHLKGMEKE-----EQEPEGLSCS 414
Db 355 YSLDKERKFLSLASNPPELL-----NLPSSVIKKVHFSG--ESKENIMRSENSESQITSK 407
QY 415 PQDED 419
Db 408 SKCKD 412

RESULT 4
US-09-173-581-6
; Sequence 6, Application US/09173581A
; Patent No. 6013455
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Tang, Y. Tcm
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina
; APPLICANT: Azimzai, Yalda
; APPLICANT: Lu, Aina
; TITLE OF INVENTION: Protein Kinase Homologs
; FILE REFERENCE: PF-0614 US
; CURRENT APPLICATION NUMBER: US/09/173,581A
```

```

; CURRENT FILING DATE: 1998-10-15
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PERL Program
; SEQ ID NO 6
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 1567782
US-09-173-581-6

Query Match
Best Local Similarity 19.7%; Score 493; DB 3; Length 345;
Matches 116; Conservative 89; Mismatches 119; Indels 96; Gaps 11;

QY 28 RYVLQKLGSGSGFTVYVSDKKAKEGELKVLKEISVGEINPNETVQANLEAQLLSKLD 87
Db 3 KYVRLQKIGESFGKAILV--KSTEDGRQY-VIKEINISRMSSKERESRREAVLANMK 59
QY 88 HPAIVKPHASVFEQDNFCIITEYCEGRDLDKIQEYKQAGKIPPENQIIEWFQILLGVD 147
Db 60 HPIVQYRESF-----EG-----ILDMFVQICLALK 85
QY 148 YMHERILHDLKSKNVFL-KNNLLKIGDFGVSRLMGSCDLATTLTGTPTHYMSPEALKH 206
Db 86 HVHDKILHRDIKSNIFLTGDTGVLGDFGIARVLNSTVELARTICIGTPYLSPEICEN 145
QY 207 QGYDTKSDIWSLACILYEMCCMHAFAGSNFLSVILKIVEGTPSLPERRYKELNAINES 266
Db 146 KPYNNKSDIHWALGCVLYELCTLKHAFAGSMKNVLKIIISGSPFVPSLHYSYDLRLSVSQ 205
QY 267 MLNKNPSLRPSAIEILKIPYLDEQLNLM-----CRYSEMTLEKNDLDCQKEAAHI 317
Db 206 LFXENPRDRPSVNSILEKGFIAKRIEFLSPOLIAEFCLTKFSKFGSQPIPAKRPSAG- 264
QY 318 INAMOKRIHLQTLRALSEVQKMT-PRERMLRKLQAADEKARKLKIVEEKYEENSKRMQ 376
Db 265 -----QNSISVMPAQKITKPAKYGI-----PLAYKKYGDKK----- 296
QY 377 ELRSRNFQQLSDVVLHEKTHLKGMEKEQEPGRSLSCSPQDEDERWQGREESDEPTLE 436
Db 297 -----LHEKKPLQKHQAHTPEKRVNTG-----EERRKISEEAAKRRLE 337

RESULT 6
US-09-992-481-2
; Sequence 2, Application US/09992481
; Patent No. 6593125
; GENERAL INFORMATION:
; APPLICANT: Fricdole, Carl Johan
; APPLICANT: Hilbun, Erin
; APPLICANT: Mathur, Brian
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. 6593125el Human Kinases and Polynucleotides
; FILE REFERENCE: Encoding the Same
; CURRENT APPLICATION NUMBER: US/09/992,481
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: US 60/252,011
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 692
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-992-481-2

Query Match
Best Local Similarity 18.2%; Score 454; DB 4; Length 692;
Matches 97; Conservative 61; Mismatches 103; Indels 8; Gaps 4;

QY 28 RYVLQKLGSGSGFTVYVSDKKAKEGELKVLKEISVGEINPNETVQANLEAQLLSKLD 87
Db 3 KYERIRVVGGAFIGVHLC---LRKADQKLVIIKQIPVEQMTKEERQAQNECQVLLN 59
QY 88 HPAIVKPHASVFEQDNFCIITEYCEGRDLDKIQEYKQAGKIPPENQIIEWFQILLGVD 147
Db 60 HPNVIEYENFLEDKALMIAMEYAPGGTIAEFIQ--KRCNSLLEESTILHFFVQILLALH 117
QY 148 YMHERILHDLKSKNVFLKNN--LLKIGDFGVSRLMGSCDLATTLTGTPTHYMSPEALK 205
Db 118 HVHDLILHRDLKTNILLDKHRNVVVGIGDFGLSKIL-SKSKAYTVVGTGTCYISPELCE 176

; CURRENT FILING DATE: 1998-10-15
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PERL Program
; SEQ ID NO 6
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 1567782
US-09-420-915-6

Query Match
Best Local Similarity 19.7%; Score 493; DB 3; Length 345;
Matches 116; Conservative 89; Mismatches 119; Indels 96; Gaps 11;

QY 28 RYVLQKLGSGSGFTVYVSDKKAKEGELKVLKEISVGEINPNETVQANLEAQLLSKLD 87
Db 3 KYVRLQKIGESFGKAILV--KSTEDGRQY-VIKEINISRMSSKERESRREAVLANMK 59
QY 88 HPAIVKPHASVFEQDNFCIITEYCEGRDLDKIQEYKQAGKIPPENQIIEWFQILLGVD 147
Db 60 HPIVQYRESF-----EG-----ILDMFVQICLALK 85
QY 148 YMHERILHDLKSKNVFL-KNNLLKIGDFGVSRLMGSCDLATTLTGTPTHYMSPEALKH 206
Db 86 HVHDKILHRDIKSNIFLTGDTGVLGDFGIARVLNSTVELARTICIGTPYLSPEICEN 145
QY 207 QGYDTKSDIWSLACILYEMCCMHAFAGSNFLSVILKIVEGTPSLPERRYKELNAINES 266
Db 146 KPYNNKSDIHWALGCVLYELCTLKHAFAGSMKNVLKIIISGSPFVPSLHYSYDLRLSVSQ 205
QY 267 MLNKNPSLRPSAIEILKIPYLDEQLNLM-----CRYSEMTLEKNDLDCQKEAAHI 317
Db 206 LFXENPRDRPSVNSILEKGFIAKRIEFLSPOLIAEFCLTKFSKFGSQPIPAKRPSAG- 264
QY 318 INAMOKRIHLQTLRALSEVQKMT-PRERMLRKLQAADEKARKLKIVEEKYEENSKRMQ 376
Db 265 -----QNSISVMPAQKITKPAKYGI-----PLAYKKYGDKK----- 296
QY 377 ELRSRNFQQLSDVVLHEKTHLKGMEKEQEPGRSLSCSPQDEDERWQGREESDEPTLE 436
Db 297 -----LHEKKPLQKHQAHTPEKRVNTG-----EERRKISEEAAKRRLE 337

RESULT 5
US-09-420-915-6
; Sequence 6, Application US/09420915
; Patent No. 6264947
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Tang, Y. Tom
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina
; APPLICANT: Azimzai, Valda
; APPLICANT: Lu, Aina
; TITLE OF INVENTION: Protein Kinase Homologs
; FILE REFERENCE: PP-0614 US
; CURRENT APPLICATION NUMBER: US/09/420,915
; CURRENT FILING DATE: 1999-10-20
; EARLIER APPLICATION NUMBER: US 09/173,581
; EARLIER FILING DATE: 1998-10-15
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PERL Program
; SEQ ID NO 6
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 1567782
US-09-420-915-6
```

QY 206 HQGYDTKSDIWSLACILYEMCCMHAFAGSNFLSIVLKIIVEGDTPLSPERYPKELNALME 265  
DB 177 GRPNQKSDIWSLACILYEMCCMHAFAGSNFLSIVLKIIVEGDTPLSPERYPKELNALME 236  
QY 266 SMLNKPNPSLRPSAIELKIPTIDELQNL 294  
DB 237 SLSLEPAQRPLSHIMAQPLCIRALLNL 265

RESULT 7  
US-09-579-664B-9  
; Sequence 9, Application US/09579664B  
; Patent No. 6514719  
; GENERAL INFORMATION:  
; APPLICANT: Immunex Corporation  
; APPLICANT: Bird, Timothy A.  
; APPLICANT: Virca, G. Duke  
; APPLICANT: Martin, Unja  
; APPLICANT: Anderson, Dirk M.  
; TITLE OF INVENTION: NOVEL MURINE AND HUMAN KINASES  
; FILE REFERENCE: 2923-A  
; CURRENT APPLICATION NUMBER: US/09/579,664B  
; CURRENT FILING DATE: 2000-05-26  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 9  
; LENGTH: 302  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-579-664B-9

Query Match 17.8%; Score 445; DB 4; Length 302;  
Best Local Similarity 36.2%; Pred. No. 2e-27;  
Matches 92; Conservative 53; Mismatches 95; Indels 14; Gaps 7;  
QY 31 LOOKLGSFGTGYLVSDKAKGEBELKV-LKEISVGL-NPNETVQANLEAQLLSKLDH 86  
DB 36 IEKKIGRGQFSEVYRAACCLDGVVVKLGDLGRFFSSKTTAAHSLVGTPTVYMSPERIHEN 89  
QY 87 DPAIVKPHASFEVDNFCIIIEYCEGRDLDDKIOEYKQAGKIFPENQIIEWFQILLGV 146  
DB 90 NHPNVKYYASFIEDNELNIVLELADAGDLGRFFSSKTTAAHSLVGTPTVYMSPERIHEN 149  
QY 147 DYMERRILHRDLKSKNVFL-KNNLLKIGDFGVSRLLMGSCDLATLTGTPTVYMSPEALK 205  
DB 150 DHMSRRVHREDIKPANVITATGVVVKLGDLGRFFSSKTTAAHSLVGTPTVYMSPERIHEN 209  
QY 206 HQGYDTKSDIWSLACILYEMCCMHAFAGSNFLSIVLKIIVEGDTPLSP-ERYPKELNA 262  
DB 210 ENGNFKSDIWSLACILYEMCCMHAFAGSNFLSIVLKIIVEGDTPLSP-ERYPKELNA 269  
QY 263 IMESMLNKNPSLRP 276  
DB 270 LVNINPDEKRP 283

RESULT 8  
US-09-221-235-2  
; Sequence 2, Application US/09221235  
; Patent No. 6043040  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: WNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,235  
; CURRENT FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE:  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 302

; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-221-235-2  
Query Match 17.7%; Score 443; DB 3; Length 302;  
Best Local Similarity 34.6%; Pred. No. 2.9e-27;  
Matches 90; Conservative 58; Mismatches 102; Indels 10; Gaps 6;  
QY 31 LOOKLGSFGTGYLVSDKAKGEBELKV-LKEISVGL-NPNETVQANLEAQLLSKLDH 88  
DB 36 IEKKIGRGQFSEVYRAACCLDGVVVKLGDLGRFFSSKTTAAHSLVGTPTVYMSPERIHEN 91  
QY 89 PAIVKPHASFEVDNFCIIIEYCEGRDLDDKIOEYKQAGKIFPENQIIEWFQILLGV 148  
DB 92 PNVIKYVASFIEDNELNIVLELADAGDLGRFFSSKTTAAHSLVGTPTVYMSPERIHEN 151  
QY 149 MHERRILHRDLKSKNVFL-KNNLLKIGDFGVSRLLMGSCDLATLTGTPTVYMSPEALKHQ 207  
DB 152 MHSRRVHREDIKPANVITATGVVVKLGDLGRFFSSKTTAAHSLVGTPTVYMSPERIHEN 211  
QY 208 GYDTKSDIWSLACILYEMCCMHAFAGSNFLSIVLKIIVEGDTPLSP-ERYPKELNALME 264  
DB 212 GYNFKSDIWSLACILYEMCCMHAFAGSNFLSIVLKIIVEGDTPLSP-ERYPKELNALME 271  
QY 265 ESMNLNKNPSLRPSAIELKI 284  
DB 272 NMCINPDEKRPDVTVYDV 291

RESULT 9  
US-09-221-928-2  
; Sequence 2, Application US/09221928  
; Patent No. 6121030  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: WNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,928  
; CURRENT FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE:  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 302  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-221-928-2

Query Match 17.7%; Score 443; DB 3; Length 302;  
Best Local Similarity 34.6%; Pred. No. 2.9e-27;  
Matches 90; Conservative 58; Mismatches 102; Indels 10; Gaps 6;  
QY 31 LOOKLGSFGTGYLVSDKAKGEBELKV-LKEISVGL-NPNETVQANLEAQLLSKLDH 88  
DB 36 IEKKIGRGQFSEVYRAACCLDGVVVKLGDLGRFFSSKTTAAHSLVGTPTVYMSPERIHEN 91  
QY 89 PAIVKPHASFEVDNFCIIIEYCEGRDLDDKIOEYKQAGKIFPENQIIEWFQILLGV 148  
DB 92 PNVIKYVASFIEDNELNIVLELADAGDLGRFFSSKTTAAHSLVGTPTVYMSPERIHEN 151  
QY 149 MHERRILHRDLKSKNVFL-KNNLLKIGDFGVSRLLMGSCDLATLTGTPTVYMSPEALKHQ 207  
DB 152 MHSRRVHREDIKPANVITATGVVVKLGDLGRFFSSKTTAAHSLVGTPTVYMSPERIHEN 211  
QY 208 GYDTKSDIWSLACILYEMCCMHAFAGSNFLSIVLKIIVEGDTPLSP-ERYPKELNALME 264  
DB 212 GYNFKSDIWSLACILYEMCCMHAFAGSNFLSIVLKIIVEGDTPLSP-ERYPKELNALME 271  
QY 265 ESMNLNKNPSLRPSAIELKI 284  
DB 272 NMCINPDEKRPDVTVYDV 291

```
RESULT 10
US-09-221-527-2
; Sequence 2, Application US/09221527
; Patent No. 6146832
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,527
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-527-2

Query Match
Best Local Similarity 17.7%; Score 443; DB 3; Length 302;
Matches 90; Conservative 58; Mismatches 102; Indels 10; Gaps 6;

QY 31 LOOKLGSFGTGYLVSDKKAKGEBELKV-LKEISVGL-NFNETVQANLEAQLSKLDH 88
Db 36 IEKKIGRGQFSEVY----RAACLLDGVVVALKKVQIFDLMDAKARADCIKEIDLKQLNH 91

QY 89 PAIVKPHASFVEQDNFCIITEYCEGRDLDKIQBYKQAGKIFPENQIIEWFIQLLGVYD 148
Db 92 PNVIKYASFIEDNELNIVLEADAGDLSRMIKHFKKQKRLIPERTVWKYFVQLCSALEH 151

QY 149 MHERRILHRLDKSKNVFL-KNNLLKIGDFGVSRLLMGSCDLATTLTGTPTHYMSPEALKHQ 207
Db 152 MHSRRVHRDIKPNVFTATGVVKGDLGLGRFFSSKTTAAHSLVGTPTHYMSPERIHEN 211

QY 208 GYDTKSDIWSLACILYEMCCMHAFAGS--NFLSIVLKIIVEGDTSPSLP-ERYPKELNAIM 264
Db 212 GYNFKSDIWSLGLLYEMAALQSPFYGDKNLYSLCKKIEQCDYPPLPFSDHYSELRQLV 271

QY 265 ESMLNKNPSLRPSAIELKI 284
Db 272 NMCINPDPEKRPDVTYVDV 291

US-09-221-527-2

Query Match
Best Local Similarity 17.7%; Score 443; DB 3; Length 302;
Matches 90; Conservative 58; Mismatches 102; Indels 10; Gaps 6;

QY 31 LOOKLGSFGTGYLVSDKKAKGEBELKV-LKEISVGL-NFNETVQANLEAQLSKLDH 88
Db 36 IEKKIGRGQFSEVY----RAACLLDGVVVALKKVQIFDLMDAKARADCIKEIDLKQLNH 91

QY 89 PAIVKPHASFVEQDNFCIITEYCEGRDLDKIQBYKQAGKIFPENQIIEWFIQLLGVYD 148
Db 92 PNVIKYASFIEDNELNIVLEADAGDLSRMIKHFKKQKRLIPERTVWKYFVQLCSALEH 151

QY 149 MHERRILHRLDKSKNVFL-KNNLLKIGDFGVSRLLMGSCDLATTLTGTPTHYMSPEALKHQ 207
Db 152 MHSRRVHRDIKPNVFTATGVVKGDLGLGRFFSSKTTAAHSLVGTPTHYMSPERIHEN 211

QY 208 GYDTKSDIWSLACILYEMCCMHAFAGS--NFLSIVLKIIVEGDTSPSLP-ERYPKELNAIM 264
Db 212 GYNFKSDIWSLGLLYEMAALQSPFYGDKNLYSLCKKIEQCDYPPLPFSDHYSELRQLV 271

QY 265 ESMLNKNPSLRPSAIELKI 284
Db 272 NMCINPDPEKRPDVTYVDV 291

US-09-221-527-2

Query Match
Best Local Similarity 17.7%; Score 443; DB 3; Length 302;
Matches 90; Conservative 58; Mismatches 102; Indels 10; Gaps 6;

QY 31 LOOKLGSFGTGYLVSDKKAKGEBELKV-LKEISVGL-NFNETVQANLEAQLSKLDH 88
Db 36 IEKKIGRGQFSEVY----RAACLLDGVVVALKKVQIFDLMDAKARADCIKEIDLKQLNH 91

QY 89 PAIVKPHASFVEQDNFCIITEYCEGRDLDKIQBYKQAGKIFPENQIIEWFIQLLGVYD 148
Db 92 PNVIKYASFIEDNELNIVLEADAGDLSRMIKHFKKQKRLIPERTVWKYFVQLCSALEH 151

QY 149 MHERRILHRLDKSKNVFL-KNNLLKIGDFGVSRLLMGSCDLATTLTGTPTHYMSPEALKHQ 207
Db 152 MHSRRVHRDIKPNVFTATGVVKGDLGLGRFFSSKTTAAHSLVGTPTHYMSPERIHEN 211

QY 208 GYDTKSDIWSLACILYEMCCMHAFAGS--NFLSIVLKIIVEGDTSPSLP-ERYPKELNAIM 264
Db 212 GYNFKSDIWSLGLLYEMAALQSPFYGDKNLYSLCKKIEQCDYPPLPFSDHYSELRQLV 271

QY 265 ESMLNKNPSLRPSAIELKI 284
Db 272 NMCINPDPEKRPDVTYVDV 291

US-09-221-527-2

Query Match
Best Local Similarity 17.7%; Score 443; DB 3; Length 302;
Matches 90; Conservative 58; Mismatches 102; Indels 10; Gaps 6;

QY 31 LOOKLGSFGTGYLVSDKKAKGEBELKV-LKEISVGL-NFNETVQANLEAQLSKLDH 88
Db 36 IEKKIGRGQFSEVY----RAACLLDGVVVALKKVQIFDLMDAKARADCIKEIDLKQLNH 91
```

```
QY 89 PAIVKPHASFVEQDNFCIITEYCEGRDLDKIQBYKQAGKIFPENQIIEWFIQLLGVYD 148
Db 92 PNVIKYASFIEDNELNIVLEADAGDLSRMIKHFKKQKRLIPERTVWKYFVQLCSALEH 151

QY 149 MHERRILHRLDKSKNVFL-KNNLLKIGDFGVSRLLMGSCDLATTLTGTPTHYMSPEALKHQ 207
Db 152 MHSRRVHRDIKPNVFTATGVVKGDLGLGRFFSSKTTAAHSLVGTPTHYMSPERIHEN 211

QY 208 GYDTKSDIWSLACILYEMCCMHAFAGS--NFLSIVLKIIVEGDTSPSLP-ERYPKELNAIM 264
Db 212 GYNFKSDIWSLGLLYEMAALQSPFYGDKNLYSLCKKIEQCDYPPLPFSDHYSELRQLV 271

QY 265 ESMLNKNPSLRPSAIELKI 284
Db 272 NMCINPDPEKRPDVTYVDV 291

RESULT 12
US-09-221-416-2
; Sequence 2, Application US/09221416
; Patent No. 6153417
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,416
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-416-2

Query Match
Best Local Similarity 17.7%; Score 443; DB 3; Length 302;
Matches 90; Conservative 58; Mismatches 102; Indels 10; Gaps 6;

QY 31 LOOKLGSFGTGYLVSDKKAKGEBELKV-LKEISVGL-NFNETVQANLEAQLSKLDH 88
Db 36 IEKKIGRGQFSEVY----RAACLLDGVVVALKKVQIFDLMDAKARADCIKEIDLKQLNH 91

QY 89 PAIVKPHASFVEQDNFCIITEYCEGRDLDKIQBYKQAGKIFPENQIIEWFIQLLGVYD 148
Db 92 PNVIKYASFIEDNELNIVLEADAGDLSRMIKHFKKQKRLIPERTVWKYFVQLCSALEH 151

QY 149 MHERRILHRLDKSKNVFL-KNNLLKIGDFGVSRLLMGSCDLATTLTGTPTHYMSPEALKHQ 207
Db 152 MHSRRVHRDIKPNVFTATGVVKGDLGLGRFFSSKTTAAHSLVGTPTHYMSPERIHEN 211

QY 208 GYDTKSDIWSLACILYEMCCMHAFAGS--NFLSIVLKIIVEGDTSPSLP-ERYPKELNAIM 264
Db 212 GYNFKSDIWSLGLLYEMAALQSPFYGDKNLYSLCKKIEQCDYPPLPFSDHYSELRQLV 271

QY 265 ESMLNKNPSLRPSAIELKI 284
Db 272 NMCINPDPEKRPDVTYVDV 291

RESULT 13
US-09-221-245-2
; Sequence 2, Application US/09221245
; Patent No. 6180358
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,245
; CURRENT FILING DATE: 1998-12-28
```

us-10-803-278-4.rai

Mon Oct 4 10:35:28 2004

EARLIER APPLICATION NUMBER: US 09/163,115  
 EARLIER FILING DATE: 1998-09-29  
 NUMBER OF SEQ ID NOS: 15  
 SOFTWARE: Patent in Ver. 2.0  
 SEQ ID NO 2  
 LENGTH: 302  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-221-245-2

Query Match 17.7%; Score 443; DB 3; Length 302;  
 Best Local Similarity 34.6%; Pred. No. 2.9e-27;  
 Matches 90; Conservative 58; Mismatches 102; Indels 10; Gaps 6;  
 QY 31 LOOKLGSFGTLYVSDKKAKGEBELKV-LKEISVGL-NPNETVQANLEAQLSKLDH 88  
 Db 36 IEKKIGRGQSEVY----RAACLDGVPVALKKQVIFDLMDAKARADCKEIDLKQLNH 91  
 QY 89 PAIVKPHASFEQDNFCIIITEYCEGRDLDDKIQEYKQAGKIPPENQIIIEWFIQLLGVDY 148  
 Db 92 PNVIKYASFIEDNELNIVLEADAGLSRMKHKFKQKRLIPERTVWKYFVQLCSALEH 151  
 QY 149 MHERRILHRLKSKNVFL-KNNLLKIGDFGVSRLLMGSCDLATTLTGTPHYMSPEALKHQ 207  
 Db 152 MHSRRVMHRDIKPNVFIATGVVKGDLGLGRFFSSKTTAAHSLVGTPTVYMSPERIHEN 211  
 QY 208 GYDTKSDIWSLACILYEMCMNHAFAGS--NFLSIVLKIVEGDTPSLP-ERYPKELNAIM 264  
 Db 212 GYNFKSDIWSLGLLYEMAALQSPFYGDKNLYSLCKIEQCDYPPPLPSDHYSEELRLQV 271  
 QY 265 ESMLNKNPSLRPSAIELKI 284  
 Db 272 NMCINPDPEKRPDVTYVDV 291

RESULT 14  
 US-09-163-115-2  
 Sequence 2, Application US/09163115A  
 Patent No. 6183962  
 GENERAL INFORMATION:  
 APPLICANT: Acton, Susan  
 TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
 FILE REFERENCE: NMI-050  
 CURRENT APPLICATION NUMBER: US/09/163,115A  
 CURRENT FILING DATE: 1998-09-29  
 NUMBER OF SEQ ID NOS: 15  
 SOFTWARE: Patent in Ver. 2.0  
 SEQ ID NO 2  
 LENGTH: 302  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-163-115-2

Query Match 17.7%; Score 443; DB 3; Length 302;  
 Best Local Similarity 34.6%; Pred. No. 2.9e-27;  
 Matches 90; Conservative 58; Mismatches 102; Indels 10; Gaps 6;  
 QY 31 LOOKLGSFGTLYVSDKKAKGEBELKV-LKEISVGL-NPNETVQANLEAQLSKLDH 88  
 Db 36 IEKKIGRGQSEVY----RAACLDGVPVALKKQVIFDLMDAKARADCKEIDLKQLNH 91  
 QY 89 PAIVKPHASFEQDNFCIIITEYCEGRDLDDKIQEYKQAGKIPPENQIIIEWFIQLLGVDY 148  
 Db 92 PNVIKYASFIEDNELNIVLEADAGLSRMKHKFKQKRLIPERTVWKYFVQLCSALEH 151  
 QY 149 MHERRILHRLKSKNVFL-KNNLLKIGDFGVSRLLMGSCDLATTLTGTPHYMSPEALKHQ 207  
 Db 152 MHSRRVMHRDIKPNVFIATGVVKGDLGLGRFFSSKTTAAHSLVGTPTVYMSPERIHEN 211  
 QY 208 GYDTKSDIWSLACILYEMCMNHAFAGS--NFLSIVLKIVEGDTPSLP-ERYPKELNAIM 264  
 Db 212 GYNFKSDIWSLGLLYEMAALQSPFYGDKNLYSLCKIEQCDYPPPLPSDHYSEELRLQV 271

QY 265 ESMLNKNPSLRPSAIELKI 284  
 Db 272 NMCINPDPEKRPDVTYVDV 291  
 RESULT 15  
 US-09-221-528-2  
 Sequence 2, Application US/09221528  
 Patent No. 6190874  
 GENERAL INFORMATION:  
 APPLICANT: Acton, Susan  
 TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
 FILE REFERENCE: NMI-050  
 CURRENT APPLICATION NUMBER: US/09/221,528  
 CURRENT FILING DATE: 1998-12-28  
 EARLIER APPLICATION NUMBER: 09/163,115  
 EARLIER FILING DATE: 1998-09-29  
 NUMBER OF SEQ ID NOS: 15  
 SOFTWARE: Patent in Ver. 2.0  
 SEQ ID NO 2  
 LENGTH: 302  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-221-528-2  
 Query Match 17.7%; Score 443; DB 3; Length 302;  
 Best Local Similarity 34.6%; Pred. No. 2.9e-27;  
 Matches 90; Conservative 58; Mismatches 102; Indels 10; Gaps 6;  
 QY 31 LOOKLGSFGTLYVSDKKAKGEBELKV-LKEISVGL-NPNETVQANLEAQLSKLDH 88  
 Db 36 IEKKIGRGQSEVY----RAACLDGVPVALKKQVIFDLMDAKARADCKEIDLKQLNH 91  
 QY 89 PAIVKPHASFEQDNFCIIITEYCEGRDLDDKIQEYKQAGKIPPENQIIIEWFIQLLGVDY 148  
 Db 92 PNVIKYASFIEDNELNIVLEADAGLSRMKHKFKQKRLIPERTVWKYFVQLCSALEH 151  
 QY 149 MHERRILHRLKSKNVFL-KNNLLKIGDFGVSRLLMGSCDLATTLTGTPHYMSPEALKHQ 207  
 Db 152 MHSRRVMHRDIKPNVFIATGVVKGDLGLGRFFSSKTTAAHSLVGTPTVYMSPERIHEN 211  
 QY 208 GYDTKSDIWSLACILYEMCMNHAFAGS--NFLSIVLKIVEGDTPSLP-ERYPKELNAIM 264  
 Db 212 GYNFKSDIWSLGLLYEMAALQSPFYGDKNLYSLCKIEQCDYPPPLPSDHYSEELRLQV 271  
 QY 265 ESMLNKNPSLRPSAIELKI 284  
 Db 272 NMCINPDPEKRPDVTYVDV 291  
 Search completed: October 1, 2004, 11:22:09  
 Job time : 34 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 11:20:51 ; Search time 135 Seconds  
(without alignments)  
1148.942 Million cell updates/sec

Title: US-10-803-278-4

Perfect score: 2501

Sequence: 1 MLKFOEAAKCVSGSTAISTY.....LGYHGDGNLISLDEYWKNEK 482

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA:\*
- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
  - 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
  - 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
  - 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
  - 12: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
  - 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
  - 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
  - 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
  - 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
  - 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
  - 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2501	100.0	482	16	US-10-803-278-4
2	2408	96.3	645	12	US-10-333-314-5
3	2408	96.3	645	16	US-10-803-278-2
4	2408	96.3	653	12	US-10-425-114-37518
5	2408	96.3	653	12	US-10-425-114-53830
6	2405	96.2	645	9	US-09-731-231A-2
7	2405	96.2	645	16	US-10-751-985-2
8	1799.5	72.0	502	12	US-10-425-114-54112
9	1491	59.6	467	10	US-09-866-050A-513
10	1480	59.2	463	16	US-09-731-231A-4
11	1480	59.2	463	16	US-10-751-985-4
12	615	24.6	841	14	US-10-198-070-87
13	615	24.6	850	12	US-10-425-114-56926
14	615	24.6	850	12	US-10-425-114-56928
15	613	24.5	841	14	US-10-198-070-52

US-10-803-278-4									
Query Match									
Best Local Similarity 100.0%; Score 2501; DB 16; Length 482;									
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MLKFOEAAKCVSGSTAI	SYPKTLTARRVVLQKLGSGSFGTVYLVSDKAKKEGELKVL	60					
Db	1	MLKFOEAAKCVSGSTAI	SYPKTLTARRVVLQKLGSGSFGTVYLVSDKAKKEGELKVL	60					
QY	61	KEISVGEINPNETVQANLEAQLLSKLDP	PAIVKFIHASFVEQDNFCIITTYCEGRDLDK	120					
Db	61	KEISVGEINPNETVQANLEAQLLSKLDP	PAIVKFIHASFVEQDNFCIITTYCEGRDLDK	120					
QY	121	QBYKQAGKIFPENQIIEWFIQLLGLVDYMHERRILHRDLKSNVFLKNNLLIKIGDFGVSR		180					

Db 121 QEQKQAGKIPFENQIIEWFIQLLGVYDVMHERRILHRDLKSNVFLKNNLLKIGDFGVS 180  
Qy 181 LLMGSCDLATLTGTGTHYMSPEALKHQGYDTKSDIWSLACILYEMCMNHAFAGSNFLSI 240  
Db 181 LLMGSCDLATLTGTGTHYMSPEALKHQGYDTKSDIWSLACILYEMCMNHAFAGSNFLSI 240  
Qy 241 VLKIVEGDTSPSPERYPKELNAIMESMLNKNPSLPSAIEILKIPYLBQQLNLMCRYSE 300  
Db 241 VLKIVEGDTSPSPERYPKELNAIMESMLNKNPSLPSAIEILKIPYLBQQLNLMCRYSE 300  
Qy 301 MTLKEDKNCQKEAAHIINAMQRIHLQTLRALSEVQKMTPRRMLRLKLOADEKARKL 360  
Db 301 MTLKEDKNCQKEAAHIINAMQRIHLQTLRALSEVQKMTPRRMLRLKLOADEKARKL 360  
Qy 361 KKIIVEKYENSKRMQELSRNFQQLSDVVLHKTGLKGMEEKEQPEGRILSCSPQDEDE 420  
Db 361 KKIIVEKYENSKRMQELSRNFQQLSDVVLHKTGLKGMEEKEQPEGRILSCSPQDEDE 420  
Qy 421 ERWQGREESDEPTLENLPESQIPSPMDLHELESIVEDATSDLYGDCNLIISLDEYKN 480  
Db 421 ERWQGREESDEPTLENLPESQIPSPMDLHELESIVEDATSDLYGDCNLIISLDEYKN 480  
Qy 481 EK 482  
Db 481 EK 482

## RESULT 2

US-10-333-314-5  
; Sequence 5, Application US/10333314  
; Publication No. US20030211093A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.; YUE, Henry  
; APPLICANT: KHAN, Farrah A.; GURURAJAN, Rajagopal  
; APPLICANT: HARFALIA, April J.A.; CHAWLA, Narinder K.  
; APPLICANT: ARVIZU, Chandra S.; RAMKUMAR, Jayalaxmi  
; APPLICANT: GANDHI, Ameena R.; POLICKY, Jennifer L.  
; APPLICANT: BAUGHN, Mariah R.; TRIBOULEY, Catherine M.  
; APPLICANT: THORNTON, Michael B.; BANDMAN, Olga  
; APPLICANT: NGUYEN, Daniel B.; LU, Yan  
; APPLICANT: BURFORD, Neil; LAL, Preeti G.  
; APPLICANT: ELLIOTT, Vicki S.; RECIPON, Shirley A.  
; APPLICANT: KEARNEY, Liam; LU, Dying Aina M.  
; APPLICANT: GREENWALD, Sara R.; TANG, Y. Tom  
; APPLICANT: XU, Yuming; WALSH, Roderick T.  
; APPLICANT: GRETZEN, Kimberly J.; YANG, Junming  
; APPLICANT: JACKSON, Jennifer L.  
; TITLE OF INVENTION: HUMAN KINASES  
; FILE REFERENCE: PI-0162 USN  
; CURRENT APPLICATION NUMBER: US/10/333,314  
; CURRENT FILING DATE: 2003-01-15  
; PRIOR APPLICATION NUMBER: PCT/US01/23092  
; PRIOR FILING DATE: 2001-07-20  
; PRIOR APPLICATION NUMBER: US 60/220,038  
; PRIOR FILING DATE: 2000-07-21  
; PRIOR APPLICATION NUMBER: US 60/222,112  
; PRIOR FILING DATE: 2000-07-28  
; PRIOR APPLICATION NUMBER: US 60/222,831  
; PRIOR FILING DATE: 2000-08-04  
; PRIOR APPLICATION NUMBER: US 60/224,729  
; PRIOR FILING DATE: 2000-08-11  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: PERL Program  
; SEQ ID NO 5  
; LENGTH: 645  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No: 7477585CD1  
US-10-333-314-5

Query Match 96.3%; Score 2408; DB 12; Length 645;  
Best Local Similarity 100.0%; Pred. No. 9,9e-157;  
Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MLKFOEAAKCVSGSTAISTYPTKTLIARRVYVLOQKLGSGSGFTVYVSDKAKRGEELKVL 60  
Db 1 MLKFOEAAKCVSGSTAISTYPTKTLIARRVYVLOQKLGSGSGFTVYVSDKAKRGEELKVL 60  
Qy 61 KEISVGLNPNETVQANLEAQLSKLDHPAIVKPHASFVEQDNFCIITEYCEGRDLDDKI 120  
Db 61 KEISVGLNPNETVQANLEAQLSKLDHPAIVKPHASFVEQDNFCIITEYCEGRDLDDKI 120  
Qy 121 QEQKQAGKIPFENQIIEWFIQLLGVYDVMHERRILHRDLKSNVFLKNNLLKIGDFGVS 180  
Db 121 QEQKQAGKIPFENQIIEWFIQLLGVYDVMHERRILHRDLKSNVFLKNNLLKIGDFGVS 180  
Qy 181 LLMGSCDLATLTGTGTHYMSPEALKHQGYDTKSDIWSLACILYEMCMNHAFAGSNFLSI 240  
Db 181 LLMGSCDLATLTGTGTHYMSPEALKHQGYDTKSDIWSLACILYEMCMNHAFAGSNFLSI 240  
Qy 241 VLKIVEGDTSPSPERYPKELNAIMESMLNKNPSLPSAIEILKIPYLBQQLNLMCRYSE 300  
Db 241 VLKIVEGDTSPSPERYPKELNAIMESMLNKNPSLPSAIEILKIPYLBQQLNLMCRYSE 300  
Qy 301 MTLKEDKNCQKEAAHIINAMQRIHLQTLRALSEVQKMTPRRMLRLKLOADEKARKL 360  
Db 301 MTLKEDKNCQKEAAHIINAMQRIHLQTLRALSEVQKMTPRRMLRLKLOADEKARKL 360  
Qy 361 KKIIVEKYENSKRMQELSRNFQQLSDVVLHKTGLKGMEEKEQPEGRILSCSPQDEDE 420  
Db 361 KKIIVEKYENSKRMQELSRNFQQLSDVVLHKTGLKGMEEKEQPEGRILSCSPQDEDE 420  
Qy 421 ERWQGREESDEPTLENLPESQIPSPMDLHELESIVEDATSDLYGDCNLIISLDEYKN 466  
Db 421 ERWQGREESDEPTLENLPESQIPSPMDLHELESIVEDATSDLYGDCNLIISLDEYKN 466

RESULT 3  
US-10-803-278-2  
; Sequence 2, Application US/10803278  
; Publication No. US20040175749A1  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Scoville, John  
; APPLICANT: Fiddie, Carl Johan  
; TITLE OF INVENTION: Novel Human Kinases and Polynucleotides Encoding the Same  
; FILE REFERENCE: LEX-0348-USA  
; CURRENT APPLICATION NUMBER: US/10/803,278  
; CURRENT FILING DATE: 2004-03-18  
; PRIOR APPLICATION NUMBER: US/10/196,927  
; PRIOR FILING DATE: 2002-05-20  
; PRIOR APPLICATION NUMBER: US 60/293,248  
; PRIOR FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 645  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-10-803-278-2

Query Match 96.3%; Score 2408; DB 16; Length 645;  
Best Local Similarity 100.0%; Pred. No. 9,9e-157;  
Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MLKFOEAAKCVSGSTAISTYPTKTLIARRVYVLOQKLGSGSGFTVYVSDKAKRGEELKVL 60  
Db 1 MLKFOEAAKCVSGSTAISTYPTKTLIARRVYVLOQKLGSGSGFTVYVSDKAKRGEELKVL 60  
Qy 61 KEISVGLNPNETVQANLEAQLSKLDHPAIVKPHASFVEQDNFCIITEYCEGRDLDDKI 120  
Db 61 KEISVGLNPNETVQANLEAQLSKLDHPAIVKPHASFVEQDNFCIITEYCEGRDLDDKI 120



QY 121 QEYKQAGKIPFENQIIEWFIQLLGVDMYHERRILHRDLKSNVFLKNLLKIGDFGYSR 180  
DB 121 QEYKQAGKIPFENQIIEWFIQLLGVDMYHERRILHRDLKSNVFLKNLLKIGDFGYSR 180  
QY 181 LLMGSCDLATTTGTGTHYMSPEALKHQGYDTKSDIWSLACILYEMCCMNHAFAGSNFLSI 240  
DB 181 LLMGSCDLATTTGTGTHYMSPEALKHQGYDTKSDIWSLACILYEMCCMNHAFAGSNFLSI 240  
QY 241 VLKIVEGDTSPSPERYPKELNAMESMLNKNPSLRPSAIELIKIPYLBOLQNLNLCRYSE 300  
DB 241 VLKIVEGDTSPSPERYPKELNAMESMLNKNPSLRPSAIELIKIPYLBOLQNLNLCRYSE 300  
QY 301 MTELEDKNLDCQKEAAHIINAMQKRIHLQTLRALSEVQKMPREMRRLKLOAADEKARKL 360  
DB 301 MTELEDKNLDCQKEAAHIINAMQKRIHLQTLRALSEVQKMPREMRRLKLOAADEKARKL 360  
QY 361 KKIIVEKYENS KRMQELSRNFQQLSVDLVHEKTHLKGMEKEEPOEGRLSCSPQDEDE 420  
DB 361 KKIIVEKYENS KRMQELSRNFQQLSVDLVHEKTHLKGMEKEEPOEGRLSCSPQDEDE 420  
QY 421 ERWQGREESDEPTLENLPESQIPSPMDLHELESIVEDATSDLGYN 466  
DB 421 ERWQGREESDEPTLENLPESQIPSPMDLHELESIVEDATSDLGYN 466

RESULT 4  
US-10-425-114-37518  
; Sequence 37518, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 37518  
; LENGTH: 653  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3547-006-G1\_FLI.pep  
US-10-425-114-37518

Query Match 96.3%; Score 2408; DB 12; Length 653;  
Best Local Similarity 100.0%; Pred. No. 1e-156;  
Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLKQEAACVSGSSTAISTYPTKLIARRVYVLOQLGSGSGFTVIVSDKAKRGEELKVL 60  
DB 9 MLKQEAACVSGSSTAISTYPTKLIARRVYVLOQLGSGSGFTVIVSDKAKRGEELKVL 68  
QY 61 KEISVGEINPNETVQANLEAQLSKLDHPAIVKPHASFVEQDNFCIITEYCEGRDLDDKI 120  
DB 69 KEISVGEINPNETVQANLEAQLSKLDHPAIVKPHASFVEQDNFCIITEYCEGRDLDDKI 128  
QY 121 QEYKQAGKIPFENQIIEWFIQLLGVDMYHERRILHRDLKSNVFLKNLLKIGDFGYSR 180  
DB 129 QEYKQAGKIPFENQIIEWFIQLLGVDMYHERRILHRDLKSNVFLKNLLKIGDFGYSR 188  
QY 181 LLMGSCDLATTTGTGTHYMSPEALKHQGYDTKSDIWSLACILYEMCCMNHAFAGSNFLSI 240  
DB 189 LLMGSCDLATTTGTGTHYMSPEALKHQGYDTKSDIWSLACILYEMCCMNHAFAGSNFLSI 248  
QY 241 VLKIVEGDTSPSPERYPKELNAMESMLNKNPSLRPSAIELIKIPYLBOLQNLNLCRYSE 300  
DB 249 VLKIVEGDTSPSPERYPKELNAMESMLNKNPSLRPSAIELIKIPYLBOLQNLNLCRYSE 308  
QY 301 MTELEDKNLDCQKEAAHIINAMQKRIHLQTLRALSEVQKMPREMRRLKLOAADEKARKL 360  
DB 309 QEYKQAGKIPFENQIIEWFIQLLGVDMYHERRILHRDLKSNVFLKNLLKIGDFGYSR 368  
QY 361 KKIIVEKYENS KRMQELSRNFQQLSVDLVHEKTHLKGMEKEEPOEGRLSCSPQDEDE 420  
DB 249 VLKIVEGDTSPSPERYPKELNAMESMLNKNPSLRPSAIELIKIPYLBOLQNLNLCRYSE 308

QY 301 MTELEDKNLDCQKEAAHIINAMQKRIHLQTLRALSEVQKMPREMRRLKLOAADEKARKL 360  
DB 309 MTELEDKNLDCQKEAAHIINAMQKRIHLQTLRALSEVQKMPREMRRLKLOAADEKARKL 368  
QY 361 KKIIVEKYENS KRMQELSRNFQQLSVDLVHEKTHLKGMEKEEPOEGRLSCSPQDEDE 420  
DB 369 KKIIVEKYENS KRMQELSRNFQQLSVDLVHEKTHLKGMEKEEPOEGRLSCSPQDEDE 428  
QY 421 ERWQGREESDEPTLENLPESQIPSPMDLHELESIVEDATSDLGYN 466  
DB 429 ERWQGREESDEPTLENLPESQIPSPMDLHELESIVEDATSDLGYN 474

RESULT 5  
US-10-425-114-53830  
; Sequence 53830, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 53830  
; LENGTH: 653  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3119-029-B6\_FLI.pep  
US-10-425-114-53830

Query Match 96.3%; Score 2408; DB 12; Length 653;  
Best Local Similarity 100.0%; Pred. No. 1e-156;  
Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLKQEAACVSGSSTAISTYPTKLIARRVYVLOQLGSGSGFTVIVSDKAKRGEELKVL 60  
DB 9 MLKQEAACVSGSSTAISTYPTKLIARRVYVLOQLGSGSGFTVIVSDKAKRGEELKVL 68  
QY 61 KEISVGEINPNETVQANLEAQLSKLDHPAIVKPHASFVEQDNFCIITEYCEGRDLDDKI 120  
DB 69 KEISVGEINPNETVQANLEAQLSKLDHPAIVKPHASFVEQDNFCIITEYCEGRDLDDKI 128  
QY 121 QEYKQAGKIPFENQIIEWFIQLLGVDMYHERRILHRDLKSNVFLKNLLKIGDFGYSR 180  
DB 129 QEYKQAGKIPFENQIIEWFIQLLGVDMYHERRILHRDLKSNVFLKNLLKIGDFGYSR 188  
QY 181 LLMGSCDLATTTGTGTHYMSPEALKHQGYDTKSDIWSLACILYEMCCMNHAFAGSNFLSI 240  
DB 189 LLMGSCDLATTTGTGTHYMSPEALKHQGYDTKSDIWSLACILYEMCCMNHAFAGSNFLSI 248  
QY 241 VLKIVEGDTSPSPERYPKELNAMESMLNKNPSLRPSAIELIKIPYLBOLQNLNLCRYSE 300  
DB 249 VLKIVEGDTSPSPERYPKELNAMESMLNKNPSLRPSAIELIKIPYLBOLQNLNLCRYSE 308  
QY 301 MTELEDKNLDCQKEAAHIINAMQKRIHLQTLRALSEVQKMPREMRRLKLOAADEKARKL 360  
DB 309 MTELEDKNLDCQKEAAHIINAMQKRIHLQTLRALSEVQKMPREMRRLKLOAADEKARKL 368  
QY 361 KKIIVEKYENS KRMQELSRNFQQLSVDLVHEKTHLKGMEKEEPOEGRLSCSPQDEDE 420  
DB 369 KKIIVEKYENS KRMQELSRNFQQLSVDLVHEKTHLKGMEKEEPOEGRLSCSPQDEDE 428  
QY 421 ERWQGREESDEPTLENLPESQIPSPMDLHELESIVEDATSDLGYN 466



QY 1 MLKFOEAAKCVSGSTAISTAYPTKTLIARRVYVLOOKLGGSGFGTVYVSDKAKRGSELKVL 60  
DB 9 MLKFOEAAKCVSGSTAISTAYPTKTLIARRVYVLOOKLGGSGFGTVYVSDKAKRGSELKVL 68  
QY 61 KEISVCELNPNETVQANLEAQLSKLDHPAIVKHFASFVEQDNFCIITEYCEGRDLDDKI 120  
DB 69 KEISVCELNPNETVQANLEAQLSKLDHPAIVKHFASFVEQDNFCIITEYCEGRDLDDKI 128  
QY 121 QYKQAGKIFPNQIIEWFIQLLGVDVYMHERRILHRDLKSNVFLKNNLLKIGDFGVSR 180  
DB 129 QYKQAGKIFPNQIIEWFIQLLGVDVYMHERRILHRDLKSNVFLKNNLLKIGDFGVSR 188  
QY 181 LMGSCDLATTGTGTPHYMSPEALKHOGYDTKSDIWSLACILYEMCCMHAFAGSNFLSI 240  
DB 189 LMGSCDLATTGTGTPHYMSPEALKHOGYDTKSDIWSLACILYEMCCMHAFAGSNFLSI 244  
QY 241 VLKIVEGDTSPSLPERYPKELNAMESMLNKNPSLRPSAIEILKIPYLDLQNLQNMCRYS 300  
DB 225 -----QELDELITLESLSQPSLPCNLNDELEPSLEGTIVDLGHY 390  
QY 301 MTELEKXNDCQKEAAHIINAMOKRIHLQTLRALSEVQKMTPRERMLRLKQAADEKARKL 360  
DB 225 -----QKRIHLQTLRALSEVQKMTPRERMLRLKQAADEKARKL 263  
QY 361 KXIVEKYBENKRMQELSRNFQOOLSVDVLHEKTHLKGMEKEKEQPEGRLLSCSPQDE 420  
DB 264 KXIVEKYBENKRMQELSRNFQOOLSVDVLHEKTHLKGMEKEKEQPEGRLLSCSPQDE 323  
QY 421 ERWQGREESDPTLENLPESQIPSMDLHELESIVEDATSDLG 466  
DB 324 ERWQGREESDPTLENLPESQIPSMDLHELESIVEDATSDLG 369

RESULT 9  
US-09-866-050A-513  
; Sequence 513, Application US/09866050A  
; Publication No. US2003004071A1  
; GENERAL INFORMATION:  
; APPLICANT: Watson, James D.  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Orrust, Rene  
; APPLICANT: Murtison, James G.  
; APPLICANT: Kumble, Krishanand D.  
; TITLE OF INVENTION: Compositions Isolated From Skin Cells  
; FILE REFERENCE: 11000.1011C4U  
; CURRENT APPLICATION NUMBER: US/09/866, 050A  
; CURRENT FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 725  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 513  
; LENGTH: 467  
; TYPE: PRT  
; ORGANISM: Rat  
US-09-866-050A-513

Query Match 59.6%; Score 1491; DB 10; Length 467;  
Best Local Similarity 64.3%; Pred. No. 5e-94;  
Matches 301; Conservative 40; Mismatches 47; Indels 80; Gaps 5;

QY 1 MLKFOEAAKCVSGSTAISTAYPTKTLIARRVYVLOOKLGGSGFGTVYVSDKAKRGSELKVL 60  
DB 1 MLKFOETAKC--GPTAIPYPTVIAIRYVLOOKLGGSGFGTVYVSDKAKRGSELKVL 58  
QY 61 KEISVCELNPNETVQANLEAQLSKLDHPAIVKHFASFVEQDNFCIITEYCEGRDLDDKI 120  
DB 59 KEISVCELNPNETVQANLEAQLSKLDHPAIVKHFASFVEQDNFCIITEYCEGRDLDDKI 118  
QY 121 QYKQAGKIFPNQIIEWFIQLLGVDVYMHERRILHRDLKSNVFLKNNLLKIGDFGVSR 180  
DB 119 QYKQAGKIFPNQIIEWFIQLLGVDVYMHERRILHRDLKSNVFLKNNLLKIGDFGVSR 178

QY 181 LMGSCDLATTGTGTPHYMSPEALKHOGYDTKSDIWSLACILYEMCCMHAFAGSNFLSI 240  
DB 179 LMGSCELATTGTGTPHYMSPEALKHOGYDTKSDIWSLACILYEMCCMHAFAGSNFLSV 238  
QY 241 VLKIVEGDTSPSLPERYPKELNAMESMLNKNPSLRPSAIEILKIPYLDLQNLQNMCRYS 300  
DB 239 VLNIVEGNTSPSLPRYPRELINIMSRMLNKNPSLRPSAIEILKIPYLDLQNLQNMCRYS 298  
QY 301 MTELEKXNDCQKEAAHIINAMOKRIHLQTLRALSEVQKMTPRERMLRLKQAADEKARK 359  
DB 299 ATLEDKNSACQKEAAHVNAI----- 320  
QY 360 LKXIVE-KYBENKRMQELSRNFQOOLSVDVLHEKTHLKGMEKEKEQPEGRLLSCSPQDE 418  
DB 321 --KITEGRRYKNNKRAKELRSQNFSGVSAHV----- 351  
QY 419 DEERWQGREESDPTLENLPESQIPSMDLHELESIVEDATSDLG 466  
DB 352 -----QELDELITLESLSQPSLPCNLNDELEPSLEGTIVDLGHY 390

## RESULT 10

US-09-731-231A-4

; Sequence 4, Application US/09731231A

; Patent No. US20020082189A1

; GENERAL INFORMATION:

; APPLICANT: GUEGLER, Karl et al

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; FILE REFERENCE: CL001007

; CURRENT APPLICATION NUMBER: US/09/731,231A

; CURRENT FILING DATE: 2000-12-07

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4

; LENGTH: 463

; TYPE: PRT

; ORGANISM: Human

US-09-731-231A-4

Query Match 59.2%; Score 1480; DB 9; Length 463;  
Best Local Similarity 100.0%; Pred. No. 2.8e-93;  
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 MGSCDLATTGTGTPHYMSPEALKHOGYDTKSDIWSLACILYEMCCMHAFAGSNFLSIVL 242  
DB 1 MGSCDLATTGTGTPHYMSPEALKHOGYDTKSDIWSLACILYEMCCMHAFAGSNFLSIVL 60  
QY 243 KIVEGDTSPSLPERYPKELNAMESMLNKNPSLRPSAIEILKIPYLDLQNLQNMCRYSMT 302  
DB 61 KIVEGDTSPSLPERYPKELNAMESMLNKNPSLRPSAIEILKIPYLDLQNLQNMCRYSMT 120  
QY 303 LEDXNDCQKEAAHIINAMOKRIHLQTLRALSEVQKMTPRERMLRLKQAADEKARKLKK 362  
DB 121 LEDXNDCQKEAAHIINAMOKRIHLQTLRALSEVQKMTPRERMLRLKQAADEKARKLKK 180  
QY 363 IVEKYBENKRMQELSRNFQOOLSVDVLHEKTHLKGMEKEKEQPEGRLLSCSPQDEDEER 422  
DB 181 IVEKYBENKRMQELSRNFQOOLSVDVLHEKTHLKGMEKEKEQPEGRLLSCSPQDEDEER 240  
QY 423 WQGREESDPTLENLPESQIPSMDLHELESIVEDATSDLG 466  
DB 241 WQGREESDPTLENLPESQIPSMDLHELESIVEDATSDLG 284

## RESULT 11

US-10-751-985-4

; Sequence 4, Application US/10751985

; Publication No. US20040126861A1

; GENERAL INFORMATION:

; APPLICANT: GUEGLER, Karl et al.

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

;; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
;; TITLE OF INVENTION: THEREOF  
;; FILE REFERENCE: CL001007CON  
;; CURRENT APPLICATION NUMBER: US/10/751,985  
;; CURRENT FILING DATE: 2004-01-07  
;; NUMBER OF SEQ ID NOS: 6  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 4  
;; LENGTH: 463  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
;; US-10-751-985-4

Query Match 59.2%; Score 1480; DB 16; Length 463;  
Best Local Similarity 100.0%; Pred. No. 2.8e-93;  
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 MGSCDLATLTGTPHYMSPEALKHQGYDTKSDIWSLACILYEMCCNNAFAGSNFLSIVL 242  
DB 1 MGSCDLATLTGTPHYMSPEALKHQGYDTKSDIWSLACILYEMCCNNAFAGSNFLSIVL 60

QY 243 KIVEGDTPSLPERYPELNAIMESMLNKNPSLRPSAIEILKIPYLDQQLNLMCRYSEMT 302  
DB 61 KIVEGDTPSLPERYPELNAIMESMLNKNPSLRPSAIEILKIPYLDQQLNLMCRYSEMT 120

QY 303 LEDKNLDCKOKEAAHINAKRIHLQTLRALSEVQKWTFRMRRLKQAADEKARKLK 362  
DB 121 LEDKNLDCKOKEAAHINAKRIHLQTLRALSEVQKWTFRMRRLKQAADEKARKLK 180

QY 363 IVEEKEYSKRMQELRSNFQOLSVDVLHKLHKGMEKEQEGRLSCSPQDEDER 422  
DB 181 IVEEKEYSKRMQELRSNFQOLSVDVLHKLHKGMEKEQEGRLSCSPQDEDER 240

QY 423 WQGREESDEPTLENPEQIPSDMLHELESIVEDATSDLG 466  
DB 241 WQGREESDEPTLENPEQIPSDMLHELESIVEDATSDLG 284

RESULT 12  
US-10-198-070-87  
; Sequence 87, Application US/10198070  
; Publication No. US20030109437A1  
; GENERAL INFORMATION:  
; APPLICANT: AVERBACK, PAUL  
; APPLICANT: GEMMELL, JACK  
; TITLE OF INVENTION: PEPTIDES EFFECTIVE IN THE TREATMENT OF TUMORS AND OTHER  
; TITLE OF INVENTION: CONDITIONS REQUIRING THE REMOVAL OR DESTRUCTION OF  
; TITLE OF INVENTION: CELLS  
; FILE REFERENCE: S9003.000008  
; CURRENT APPLICATION NUMBER: US/10/198,070  
; CURRENT FILING DATE: 2002-07-19  
; PRIOR APPLICATION NUMBER: 60/306,161  
; PRIOR FILING DATE: 2001-07-19  
; PRIOR APPLICATION NUMBER: 60/306,150  
; PRIOR FILING DATE: 2001-07-19  
; PRIOR APPLICATION NUMBER: 60/331,477  
; PRIOR FILING DATE: 2001-11-16  
; NUMBER OF SEQ ID NOS: 125  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 87  
; LENGTH: 841  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-10-198-070-87

Query Match 24.6%; Score 615; DB 14; Length 841;  
Best Local Similarity 34.7%; Pred. No. 1.1e-33;  
Matches 157; Conservative 87; Mismatches 152; Indels 56; Gaps 17;

QY 29 YVLOQKLGSGSFGTVYLVSDKKAKRGEELKVLKIEISVGEINPNETVQANLEAQLSKLDH 88  
DB 6 YCYLRVVGKSGSYGEVTLV---KHRRDGKQYVIKKLNLRNASSRERRAARAEQAEQLSKLKH 62

QY 89 PAIVKPHASFEVDNFC-IITEYCEGRDLDDKIQEYKQAGKIPFENQIIIEWFIQLLLGVD 147  
DB 63 PNIVTYKESWEGGDLIYVWGCCEGDLRYRKLKE--QKQQLLPENQVVFQIAMAQ 120

QY 148 YMHERRILHRDLKSKNVFL-KNNLLKIGDGVSRLLMGSCDLATLTGTPHYMSPEALKH 206  
DB 121 YLHEKHILHRDLKTONVFLTRTNIIVKVDLGIARVLENHCDMASTLTIGTPYMSPELFSN 180

QY 207 QGYTSDIWSLACILYEMCCNNAFAGSNFLSIVLKIIVEGDTPSLPERYPELNAIMES 266  
DB 181 KPNYKSDVWALGCCVYEMATLKHAFAKDNMNSLVYRIIEGKLPAMPDRDYSPELAEIIRT 240

QY 267 MLNKNPSLRPSAIEILKIPYLDQQLNLMCRYSEMTLED--KNLDCQ-KEAAHIN---- 319  
DB 241 MLSKRPERSVRSILRQPIKQISFFLEATKITSKNNIKNGDSQSKPFAVWSGEAE 300

QY 320 AMQKRHLQTLRA-----LSEVQKWTFRMRRLKJ--QAADEKARKLKIVBEKYEN 371  
DB 301 SNHEVIHPQLSSEGSQTYIMGEGKCLSQEKPRASGLKSPASKAHTCKQDL-----SN 355

QY 372 SKRMQELRSNFQOLSVDVLHKLHKGMEE-----KBEQPE-----GRLSCSPQD 417  
DB 356 ITELATISSVN-----IDL-----PAKGRDSVSDGFVQENQPRYLDASNELGSI-CSISQ 405

QY 418 EDERWQGREESDEPTLENPEQIPSDML 449  
DB 406 VEEMLQDNTKSAQP--ENL---IPMWSSDI 432

RESULT 13  
US-10-425-114-56926  
; Sequence 56926, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovacic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 56926  
; LENGTH: 850  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURES:  
; OTHER INFORMATION: Clone ID: LIB4119-128-E10\_FLI.pep  
; US-10-425-114-56926

Query Match 24.6%; Score 615; DB 12; Length 850;  
Best Local Similarity 34.7%; Pred. No. 1.2e-33;  
Matches 157; Conservative 87; Mismatches 152; Indels 56; Gaps 17;

QY 29 YVLOQKLGSGSFGTVYLVSDKKAKRGEELKVLKIEISVGEINPNETVQANLEAQLSKLDH 88  
DB 61 YCYLRVVGKSGSYGEVTLV---KHRRDGKQYVIKKLNLRNASSRERRAARAEQAEQLSKLKH 117

QY 89 PAIVKPHASFEVDNFC-IITEYCEGRDLDDKIQEYKQAGKIPFENQIIIEWFIQLLLGVD 147  
DB 118 PNIVTYKESWEGGDLIYVWGCCEGDLRYRKLKE--QKQQLLPENQVVFQIAMAQ 175

QY 148 YMHERRILHRDLKSKNVFL-KNNLLKIGDGVSRLLMGSCDLATLTGTPHYMSPEALKH 206  
DB 176 YLHEKHILHRDLKTONVFLTRTNIIVKVDLGIARVLENHCDMASTLTIGTPYMSPELFSN 235

QY 207 QGYTSDIWSLACILYEMCCNNAFAGSNFLSIVLKIIVEGDTPSLPERYPELNAIMES 266  
DB 236 KPNYKSDVWALGCCVYEMATLKHAFAKDNMNSLVYRIIEGKLPAMPDRDYSPELAEIIRT 295

```

QY 267 MLKNFSLRPSAIEILKPYLDEQLQNLWCRIYSEMTLED--KNLDCQ-KEAAHIIN----- 319
Db 296 MLSKREPEERPSVRSILROPIKQISPFLEATKINTKNNIKNGDSQKPFATWSGEAE 355
QY 320 AMOKRIHLQTLRA-----LSEVQKWTFRMRRLKQ--QAADEKARLKKIIVEEKYEN 371
Db 356 SNHEVHPDPLSGESQTYINGEGKLSQEKPRASGLKSPASLKAHTCKQDL-----SN 410
QY 372 SKRMQELRGRNPFQQLSVDVHLKTHLKGME-----KEQPE-----GRLSGSPQD 417
Db 411 TTLETATISSVN-----IDIL-----PAKGRDSVDSQFVQENQPRYILDASNELGCI-CSISQ 460
QY 418 EDERERWQGREESDEPTLENLPESQIPSMDL 449
Db 461 VBEMLQDNTKSSAQF--ENL---IPMWSSDI 487

RESULT 14
US-10-425-114-56928
; Sequence 56928, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Zhou, Yihua
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 56928
; LENGTH: 850
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4119-128-G9_FLI.pgp
US-10-425-114-56928

```

us-10-803-278-4.rapb

Mon Oct 4 10:35:29 2004

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 11:05:15 ; Search time 25 Seconds  
(without alignments)  
1003.913 Million cell updates/sec

Title: US-10-803-278-4

Perfect score: 2501

Sequence: 1 MLKFEAAKCVSGSTAISTY.....LGVHGDGNLISLDEYWKNEK 482

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	615	24.6	841	1 NEK4 HUMAN	P51957 homo sapien
2	600.5	24.0	792	1 NEK4 MOUSE	Q921j2 mus musculus
3	598	23.9	774	1 NEK1 MOUSE	P51954 mus musculus
4	585	23.4	506	1 NEK3 HUMAN	P51956 homo sapien
5	584.5	23.4	511	1 NEK3 MOUSE	Q9r0a5 mus musculus
6	582.5	23.3	1258	1 NEK1 HUMAN	Q96py6 homo sapien
7	495	19.8	445	1 NEK2 HUMAN	P51955 homo sapien
8	474.5	19.0	697	1 NEK2 BRARE	P22209 brachydanio
9	463.5	18.5	435	1 KIN3 YEAST	Q90kc2 saccharomyc
10	454	18.2	692	1 NEK8 HUMAN	Q86596 homo sapien
11	454	18.2	698	1 NEK8 MOUSE	Q92zr4 mus musculus
12	452	18.1	443	1 NEK2 MOUSE	O35942 mus musculus
13	448	17.9	722	1 FIN1 SCHPO	O13839 schizosacch
14	445	17.8	302	1 NEK7 MOUSE	Q8es74 mus musculus
15	443	17.7	302	1 NEK7 HUMAN	Q8tdx7 homo sapien
16	443	17.7	313	1 NEK6 MOUSE	Q9es70 mus musculus
17	442	17.7	313	1 NEK6 RAT	P59895 rattus norv
18	440.5	17.6	699	1 NIMA EMENI	P11837 emericellia
19	439	17.6	313	1 NEK6 HUMAN	Q9hc98 homo sapien
20	429.5	17.2	979	1 NEK9 HUMAN	Q8td19 homo sapien
21	425.5	17.0	984	1 NEK9 MOUSE	Q8klr7 mus musculus
22	414.5	16.6	944	1 NEK9 XENLA	Q7zxc8 xenopus lae
23	410.5	16.4	779	1 NIM1 NEUCR	P48479 neurospora
24	409.5	16.4	431	1 NRKA TRYBB	Q08942 trypanosoma
25	403.5	16.1	431	1 NRKE TRYBB	Q03428 trypanosoma
26	388	15.5	446	1 KCC1 YEAST	P27466 saccharomyc
27	382.5	15.3	1062	1 CC7 SCHPO	P41892 schizosacch
28	380.5	15.2	682	1 SNK RAT	Q9r012 rattus norv
29	376.5	15.1	443	1 STJ4 HUMAN	Q9v6e0 homo sapien
30	372.5	14.9	682	1 SNK MOUSE	P53351 mus musculus
31	370.5	14.8	685	1 SNK HUMAN	Q9zyy3 homo sapien
32	368.5	14.7	447	1 KCC2 YEAST	P22517 saccharomyc
33	360.5	14.4	543	1 CHK2 HUMAN	O96017 homo sapien

ALIGNMENTS

RESULT 1

ID	NEK4 HUMAN	STANDARD;	PRT;	841 AA.
AC	P51957;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Serine/threonine-protein kinase Nek4 (EC 2.7.1.37) (NimaA-related protein kinase 4) (Serine/threonine protein kinase 2)			
DE	(Serine/threonine-protein kinase NRK2).			
GN	NEK4 OR STK2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Breast;			
EX	MEDLINE=94268838; PubMed=8208544;			
RA	Levedakou E.N., He M., Baptist E.W., Craven R.J., Cance W.G.,			
RA	Welch P.D., Simmons A., Naylor S.L., Leach R.J., Lewis T.B.,			
RA	Bowcock A., Liu E.T.;			
RT	"Two novel human serine/threonine kinases with homologies to the cell cycle regulating Xenopus MO15, and NIMA kinases: cloning and characterization of their expression pattern.";			
RT	Oncogene 9:1977-1988(1994).			
RL				
CC	-!- FUNCTION: Seems to act exclusively upon threonine residues (By similarity).			
CC	-!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.			
CC	-!- SUBCELLULAR LOCATION: Nuclear (Probable).			
CC	-!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN ADULT HEART, FOLLOWED BY PANCREAS, SKELETAL MUSCLE, BRAIN, LIVER, KIDNEY, LUNG AND PLACENTA. PRESENT IN MOST PRIMARY CARCINOMAS.			
CC	-!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIMA subfamily.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL; L20321; AAA36658.1; --			
DR	PIR; I78885; I78885.			
DR	HSSP; Q63450; I806.			
DR	Genew; HGNC:11399; NEK4.			
DR	GM; P51957; --			
DR	MM; 601959; --			
DR	GO; GO:0004674; P:protein serine/threonine kinase activity; TAS.			
DR	GO; GO:0006468; P:protein amino acid phosphorylation; TAS.			
DR	InterPro; IPR000719; Prot_kinase.			
DR	InterPro; IPR008271; Ser_Thr_pkin_AS.			
DR	InterPro; IPR002290; Ser_Chr_pkinase.			
DR	InterPro; IPR001245; Tyr_pkinase.			

P41951 caenorhabdi  
P53350 homo sapien  
Q9uk85 homo sapien  
Q94804 homo sapien  
Q95819 homo sapien  
O55098 mus musculu  
P23647 drosophila  
Q92865 mus musculu  
P97820 mus musculu  
O62673 rattus norv  
O00506 homo sapien  
P32562 saccharomyc

34 355.5 14.2 1576 1 YLK3 CAEEL  
35 354 14.2 603 1 PLK1\_HUMAN  
36 352.5 14.1 1360 1 TNIK\_HUMAN  
37 352 14.1 968 1 STKA\_HUMAN  
38 349.5 14.0 1239 1 M4K1\_HUMAN  
39 349 14.0 966 1 STKA\_MOUSE  
40 346.5 13.9 805 1 FUSE\_DROME  
41 345 13.8 546 1 CHK2\_MOUSE  
42 345 13.8 1233 1 M4K1\_MOUSE  
43 344 13.8 603 1 PLK1\_RAT  
44 342 13.7 426 1 ST25\_HUMAN  
45 342 13.7 705 1 CDC5\_YEAST

DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot kinase; 1.  
 DR SMART; SM00220; S TKC\_1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.  
 DR PROSITE; PS00111; PROTEIN KINASE ST; 1.  
 KW Transferase; Serine/threonine-protein kinase; ATP-binding; Mitosis;  
 KW Nuclear protein; Phosphorylation.  
 FT DOMAIN 6 261 PROTEIN KINASE.  
 FT NP\_BIND 12 20 ATP (BY SIMILARITY).  
 FT BINDING 35 35 ATP (BY SIMILARITY).  
 FT ACT\_SITE 131 131 BY SIMILARITY.  
 FT MOD\_RES 165 165 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 SQ SEQUENCE 841 AA; 94571 MW; 0DD31920DDE7EA58 CRC64;  
 Query Match 24.6%; Score 615; DB 1; Length 841;  
 Best Local Similarity 34.7%; Pred. No. 1e-24;  
 Matches 157; Conservative 87; Mismatches 152; Indels 56; Gaps 17;  
 QY 29 YVLQKLGSGSGFTVYLVSDKXKAGEELKVLKISVGLNPNETVQANLEAQLLSKLDH 88  
 Db 6 YCVLRVVGKSGYGEVTLV---KRRDQKQYVTKLNLNASSRRERRAAEQEAQLLSQLKH 62  
 QY 89 PAIVKHFASFEVDNFC-IITEYCEGRDLDKIQEYKQAGKIFPENQIIEWFIQLLQVD 147  
 Db 63 PNIVTYKESWEGDGLLYVMGFCGGDLYRLKE--QKQQLPENQVVFVQIAMAALQ 120  
 QY 148 YHMERILRLDLKSNVFL--KNLKLKIGDFGVSRLMGSCDLATLTGTTPHYMSPEALKH 206  
 Db 121 YLHKEHILRLDLKSNVFLTRNLIKVGDLGIAVLNHCNDVASTLIGTPHYMSPELFSN 180  
 QY 207 QGYDTHKSDIWSLACILYECMMNHAFAGSNFLSIVLKVSGTPTSLPERYPKELNAIVES 266  
 Db 181 KPNYKSDYVWALGCCVYEWATLKHAFNAKMDMSLYRIEGLKLPAMPDYSPELAEIIRT 240  
 QY 267 MLNKNPSLRPSAIEILKIPYLDQQLNLMCRYSEMLED--KNLDCQ-KEAAHIIN--- 319  
 Db 241 MLSKREPERPSVRSILRQYIKRQISFFLEATKIKTKNNKNGDSQSKPATVTVSGAE 300  
 QY 320 AMQRIHLQTLA-----LSEVQKMPERRMLRLK--QAADKARKLKIIVEEKYEN 371  
 Db 301 SNHEVTHPQLSSEGSQTYINGEKLCSQEKPRASGLLKSPLAKHAKTCQDL-----SN 355  
 QY 372 SKRMQBLRSRNFQQLSDVYLHKEHLKGMEE-----XEEQPE-----GRLSQSPQD 417  
 Db 356 TTELATISSV-----IDL-----PAKGRDSVSGFVQENQRYLDASNELGGI-CSISQ 405  
 QY 418 EDERWQGRBESDEFTLENLPESQIPSMOL 449  
 Db 406 VEEMLQDNTKSAQP--ENL---IPMWSSDI 432

## RESULT 2

NEK4\_MOUSE STANDARD; PRT; 792 AA.  
 AC Q92IJ2; O35673; Q9RIJ1;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Serine/threonine-protein kinase Nek4 (EC 2.7.1.37) (Nima-related  
 DE protein kinase 4) (Serine/threonine protein kinase 2).  
 GN NEK4 OR STK2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]\_TaxID=10090;  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
 RC STRAIN=BALB/c; TISSUE=Spleen;  
 RX MEDLINE=20001940; PubMed=10529384;  
 RA Hayashi K., Igarashi H., Ogawa M., Sakaguchi N.;  
 RT "Activity and substrate specificity of the murine STK2

RT serine/threonine kinase that is structurally related to the mitotic  
 RT regulator protein NIMA of *Aspergillus nidulans*.";  
 RL Biochem. Biophys. Res. Commun. 264:449-456(1999).  
 RN [2].  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RX MEDLINE=9321807; PubMed=10393247;  
 RA Chen A., Yanai A., Arana E., Kilfin G., Metro B.;  
 RT "NIMA-related kinases: isolation and characterization of murine nek3  
 RT and nek4 cDNAs, and chromosomal localization of nek1, nek2 and nek3.";  
 RL Gene 234:127-137(1999).  
 CC -!- FUNCTION: Seems to act exclusively upon threonine residues.  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1; Synonyms=mSTK2L;  
 CC IsoId=Q92IJ2-1; Sequence=Displayed;  
 CC Name=2; Synonyms=mSTK2S;  
 CC IsoId=Q92IJ2-2; Sequence=VSP\_007001;  
 CC and is upregulated in the testis.  
 CC -!- TISSUE SPECIFICITY: Expressed ubiquitously among various organs  
 CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIMA  
 CC subfamily.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; AJ223071; CAAL1072.1; -.  
 DR EMBL; Y09234; CAA70436.1; -.  
 DR EMBL; AF099067; AAD16287.1; -.  
 DR PIR; JC7122; JC7122.  
 DR HSP; Q63450; IAO6.  
 DR MGD; MGI:1344404; Nek4.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR08271; Ser\_Thr\_kinase.  
 DR InterPro; IPR002290; Ser\_Thr\_kinase.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot kinase; 1.  
 DR SMART; SM00220; S TKC\_1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00111; PROTEIN KINASE ST; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.  
 KW Transferase; Serine/threonine-protein kinase; ATP-binding; Mitosis;  
 KW Nuclear protein; Phosphorylation; Alternative splicing.  
 FT DOMAIN 6 261 PROTEIN KINASE.  
 FT NP\_BIND 12 20 ATP (BY SIMILARITY).  
 FT BINDING 35 35 ATP (BY SIMILARITY).  
 FT ACT\_SITE 131 131 BY SIMILARITY.  
 FT MOD\_RES 165 165 Missing (in isoform 2).  
 FT VARSPLIC 456 503 /FTID=VSP\_007001.  
 FT FT CONFLICT 499 499 G -> R (IN REF. 2).  
 SQ SEQUENCE 792 AA; 88994 MW; CF9187311C807A1D CRC64;  
 Query Match 24.0%; Score 600.5; DB 1; Length 792;  
 Best Local Similarity 31.5%; Pred. No. 5.4e-24;  
 Matches 141; Conservative 90; Mismatches 143; Indels 73; Gaps 13;  
 QY 29 YVLQKLGSGSGFTVYLVSDKXKAGEELKVLKISVGLNPNETVQANLEAQLLSKLDH 88  
 Db 6 YCVLRVVGKSGYGEVTLV---KRRDQKQYVTKLNLNASSRRERRAAEQEAQLLSQLKH 62  
 QY 89 PAIVKHFASFEVDNFC-IITEYCEGRDLDKIQEYKQAGKIFPENQIIEWFIQLLQVD 147  
 Db 63 PNIVTYKESWEGDGLLYVMGFCGGDLYRLKE--QKQQLPENQVVFVQIAMAALQ 120



QY 148 YMHERRILHRLDKSKNVFL-KNNLLKIGDFGVSRLLMGSCDLATTLTGTPHYMSPEALKH 206  
 Db 121 YLHEKHILHRLDKTQNVFLTRNIIVKVGDLGIARVLENGHGVASTLIGTPHYMSPELFSN 180  
 QY 207 QGYDTKSDIWSLACILYEMCMNHAFAGSNFLSIVLKIVGEGTDSLPRYPKELNAMES 266  
 Db 181 KPNYKSDYVALGCCVYEWATLKHAFNAKDMNSLVRIIEGKLPPMPKVIYSELAEELRT 240  
 QY 267 MLNKNPSLRPSAIEILKIPYLDEQL-----QNLRCRYSEMFL-----ED 305  
 Db 241 MLSRREPERPSVRSILRQYIKHHISLFLEATKAKTSKNNVNCDSRAKVPAAVSRKEE 300  
 QY 306 KULD-----CKEAAHIIN-----AMQKIHLOTLRALSEVQKMTPRERMRURKIQ 352  
 Db 301 SNTDVIHYQPRSESGSALVMGDKLSQEKFPDGLPLASPLASLEGHTGKQDM-----NN 355  
 QY 353 ADEKARKLAKI-----VEEKEENSCKRMQELSRNFQSLVDVLEHKLKMGEEKREQ 406  
 Db 356 TGESCATISRINDILPAERDSANAGVQE-----SQQHVDAEDVDSQ 401  
 QY 407 PEGRLSCPDDEERWQGBESDEP 433  
 Db 402 -----CSISQE-KERLQGNTRSSDQ 421

RESULT 3

NEKI\_MOUSE  
 ID NEKI\_MOUSE STANDARD; PRT; 774 AA.  
 AC P51954;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Serine/threonine-protein kinase Nek1 (EC 2.7.1.37) (NimaA-related  
 DE protein kinase 1).  
 GN NEK1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC TISSUE=Blood;  
 RA MEDLINE=93010942; PubMed=1382974;  
 RA Letwin K., Mizzen L., Motro B., Ben-David Y., Bernstein A.,  
 RA Pawson T.;  
 RT "A mammalian dual specificity protein kinase, Nek1, is related to the  
 RT Nima cell cycle regulator and highly expressed in meiotic germ  
 RT cells.";  
 RL EMBO J. 11:3521-3531(1992).  
 CC -!- FUNCTION: PHOSPHORYLATES SERINES AND THREONINES, BUT ALSO APPEARS  
 CC -!- TO POSSESS TYROSINE KINASE ACTIVITY. IMPLICATED IN THE CONTROL OF  
 CC MEIOSIS.  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -!- TISSUE SPECIFICITY: PREDOMINANTLY IN TESTES (GERM CELLS AND  
 CC SERTOLI CELLS). LOWER LEVELS IN OVARY (OVOCYTES AND GRANULOSA  
 CC CELLS), THYMUS, AND LUNG.  
 CC -!- DEVELOPMENTAL STAGE: IN FEMALE, EXPRESSED AS FOLLICLES ENTER THE  
 CC SECONDARY STAGE UNTIL OVULATION OCCURS. IN THE MALE REPRODUCTIVE  
 CC SYSTEM, THE EXPRESSION IS LIMITED TO SPERMATOCYTES AND SPERMATIDS.  
 CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIMA  
 CC subfamily.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 CC EMBL; S45828; AAB23529.1; -  
 CC PIR; S25284; S25284.

HSSP; Q63450; 1A06.  
 MGD; MGI:97303; Nek1.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR008271; Ser\_thr\_kin\_AS.  
 DR InterPro; IPR002290; Ser\_thr\_kinase.  
 DR InterPro; IPR001245; Tyr\_kinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SM00220; S\_TKC\_1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
 DR Transferase; Serine/threonine-protein kinase; ATP-binding; Mitosis;  
 KW Nuclear protein; Phosphorylation; Cell cycle; Cell division;  
 KW Tyrosine-protein kinase.  
 FT DOMAIN 4 258 PROTEIN\_KINASE.  
 FT NP\_BIND 10 18 ATP (BY SIMILARITY).  
 FT BINDING 33 33 ATP (BY SIMILARITY).  
 FT ACT\_SITE 128 128 BY SIMILARITY.  
 FT MOD\_RES 162 162 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 SQ SEQUENCE 774 AA; 88427 MW; 7EDE5881ACDC06FA CRC64;

Query Match 23.9%; Score 598; DB 1; Length 774;

Best Local Similarity 28.5%; Pred. No. 7.1e-24;  
 Matches 141; Conservative 94; Mismatches 132; Indels 128; Gaps 13;

QY 28 RYVLQKLGSGSGFTYVYVSDKKAQGEELVKLEISVGEINPNVTQANLEAQLLSKLD 87  
 Db 3 KYVRLQKIGSGSGKAVLV--KSTEDGRHY-VIKEINISRMDSKQERESRREAVLANMK 59  
 QY 88 HPAIVKSHAFVEQDNFCIITEYCEGDDDDKQYKQAGKIPENQIIEFWQLLLGVD 147  
 Db 60 HPNTVQYKESFEENGSLYIVNDYCEGGDLFKRIN--AQKALFQEDQILDWFVQICLALK 117  
 QY 148 YMHERRILHRLDKSKNVFL-KNNLLKIGDFGVSRLLMGSCDLATTLTGTPHYMSPEALKH 206  
 Db 118 YVHDKILHRLDIKSONIFLKDCTVQLGDFGIARVLENGHGVASTLIGTPHYLSPEICEN 177  
 QY 207 QGYDTKSDIWSLACILYEMCMNHAFAGSNFLSIVLKIVGEGTDSLPRYPKELNAMES 266  
 Db 178 KPNYKSDIWSLACILYEMCMNHAFAGSNFLSIVLKIVGEGTDSLPRYPKELNAMES 237  
 QY 267 MLNKNPSLRPSAIEILKIPYLDEQLNLM-----C----- 296  
 Db 238 LFRKNPRDRFSVNSILEKGFIAKRIEFLSPQIAEFCULTKSLKFGQPLPKRPSAQ 297  
 QY 297 -----RYSEMTLBDKNLDCQEAHIIINAMQKRIHLQTLR 331  
 Db 298 GVSFVPAQKITPAKYGVPLATYKYDKKLEKKPPPKHQAQI----- 344  
 QY 332 ALSVQVQVTPRMRRLKLADEKARKLKIIVEEYKENSKEMLSRNFQSLVDVL 391  
 Db 345 ---PVKQNSGEERKKMSEEA--KKRLEFIEKEK-----KQDKQIRFLKAEQMK---R 391  
 QY 392 HEKTHLKGMEKEBEQP-----EGRLS---CSQ----- 416  
 Db 392 QEQRLERINRAREQGRNVLRAGSGEVKASFFGIGGAVSPSPSPRGQYEHVHAFDQ 451  
 QY 417 -----DEDEERWQ 425  
 Db 452 MQLRAEDNEARWKG 466

RESULT 4

NEK3\_HUMAN  
 ID NEK3\_HUMAN STANDARD; PRT; 506 AA.  
 AC P51956; Q8WUN5;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Serine/threonine-protein kinase Nek3 (EC 2.7.1.37) (NimaA-related  
 DE protein kinase 3) (HSPK 36).



GN NEK3.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99240743; PubMed=10224116;  
 RA Tanaka K., Nigg E.A.;  
 RT "Cloning and characterization of the murine Nek3 protein kinase, a  
 RT novel member of the NIMA family of putative cell cycle regulators.";  
 RL J. Biol. Chem. 274:13491-13497(1999).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99321807; PubMed=10393247;  
 RA Chen A., Imani A., Arama E., Killin G., Motro B.;  
 RT "NIMA-related kinases: isolation and characterization of murine nek3  
 RT and nek4 cDNAs, and chromosomal localization of nek1, nek2 and nek3.";  
 RL Gene 234:127-137(1999).  
 CC -!- FUNCTION: Kinase that may play a role in mitotic regulation.  
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein. NIMA  
 CC -!- SIMILARITY: Belongs to the Ser/thr family of protein kinases. NIMA  
 CC subfamily.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 -----  
 DR EMBL; AF093416; AAD20986.1; -  
 DR EMBL; AF099066; AAD16286.1; -  
 DR HSP: Q63450; IA06  
 DR MGD; MGI:1344371; Nek3.  
 DR InterPro; IPR000719; Prot\_kinase.  
 DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot\_kinase; 1.  
 DR SMART; SMC0220; S\_TKC; 1.  
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.  
 DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.  
 DR TransFam; Serine/threonine-protein kinase; ATP-binding; Mitosis;  
 KW Nuclear protein; Phosphorylation; Cell cycle; Cell division.  
 FT DOMAIN 4 255 PROTEIN KINASE.  
 FT NP\_BIND 10 18 ATP (BY SIMILARITY).  
 FT BINDING 33 33 ATP (BY SIMILARITY).  
 FT ACT\_SITE 125 125 BY SIMILARITY.  
 FT MOD\_RES 159 159 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT CONFLICT 239 239 N -> K (IN REF. 2).  
 FT CONFLICT 342 343 MISSING (IN REF. 2).  
 SQ SEQUENCE 511 AA; 57222 MW; D6D6C0533C7302F CRC64;  
 -----  
 Query Match 23.4%; Score 584.5; DB 1; Length 511;  
 Best Local Similarity 34.9%; Pred. No. 2.2e-23;  
 Matches 149; Conservative  
 QY 29 YVLQKLGSGFTGYLVISDKAKRG---BELKVLKEISVGEINPNETVQANLEAQLISK 85  
 Db 4 YTVLRVIGQSGFGRALLVLQESNQTFAMKEIRLLK-----SDQTSRKEAVLLAK 54  
 QY 86 LDHPAIVKFFHASFVEQDNFCITTEYCEGRDLDDKIQBYKAGKIFPENQIIEWFIQLLLG 145  
 Db 55 MKHENIVAPKESFEAEGLYIVMVEYCDGGLMQRIKQ--QKGNLFPEDTILNWFQICLG 112  
 QY 146 VDYWHERRIILRLDKSNVFLKNN-LLKIGDFGVSRLLMGSCDLATLTGTPTMGPEAL 204  
 Db 113 VNIHKRVLHRDIKSNVFLTHNGKVLGDFGSRLLSSPMFACTYGTGYVPEPEIW 172

QY 205 KHGYDTKSDIWSLACILYEMCCMNHAFAGSNFLSIVLKIVEGDTFSLPFRYPKELNIM 264  
 :  
 Db 173 ENLPYNKNSDIWSLGCIYLCELCAKHFPQANSWKNLLIKICQGPIHPLPALYSCKLGLIV 232  
 :  
 QY 265 ESMLNKPSLRPPSAIEILKIPYLDEQLONLRCRYSEMTELDKNLDCKEAAHIINAMQKR 324  
 :  
 Db 233 KOMLRPNRSHRPSA-----TTLCRGSLAPLVPKCLPPQ-----IIREYGEQ 274  
 :  
 QY 325 IHLQTLRALSEVQKWTPRRMRLEKLQAADKARKLKIVKEKYEENSKRMQ-ELRSNF 383  
 :  
 Db 275 I-----LDEIKISTPKN---MKQDSNRVGRALG-----EANSAAQEEERGRKC 316  
 :  
 QY 384 QQLSVDVLHEKT-----HLKMWEKEQEPEGRKLSCPQEDDERVQGREESDE 432  
 :  
 Db 317 SHTEL-----ESTGTTPAGNALGRAARGNPESGNRQEHGHSHTSPASPHRPWRHERGPSNV 372  
 :  
 QY 433 PTLE 436  
 ||  
 Db 373 EALE 376  
 ||

RESULT 6

NEKI HUMAN STANDARD; PRT; 1258 AA.

ID NEKI\_HUMAN AC Q96PY6; O9Y594;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Serine/threonine-protein kinase Nek1 (EC 2.7.1.37) (Nima-related protein kinase 1) (NY-REN-55 antigen).  
 DE protein kinase 1 (NY-REN-55 antigen).  
 GN NEKI OR KIAA1901.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_Taxid=9606;  
 [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Brain;  
 RX MEDLINE=21456161; PubMed=11572484;  
 RA Nagase T., Kikuno R., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XXI. The complete sequences of 60 new cDNA clones from brain which code for large proteins";  
 RL DNA Res. 8:179-187(2001).  
 RN [2]  
 RP SEQUENCE OF 444-1258 FROM N.A. (ISOFORM 2).  
 RC TISSUE=Renal cell carcinoma;  
 RX MEDLINE=93438124; PubMed=10508479;  
 RA Scanlan M.J., Gordan J.D., Williamson B., Stockert E., Bander N.H., Jongeneel C.V., Gure A.O., Jager D., Jager E., Knuth A., Chen Y.-T., Old L.J.;  
 RA Antigen recognized by autologous antibody in patients with renal-cell carcinoma.";  
 RT Int. J. Cancer 83:456-464(1999).  
 RL  
 CC -/- FUNCTION: PHOSPHORYLATES SERINES AND THREONINES, BUT ALSO APPEARS TO POSSESS TYROSINE KINASE ACTIVITY. IMPLICATED IN THE CONTROL OF MEIOSIS (By similarity).  
 CC -/- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
 CC -/- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -/- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=Q96PY6-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q96PY6-2; Sequence=VSP\_004870;  
 CC Note=No experimental confirmation available;  
 CC -/- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIMA subfamily.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use.

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; AB067488; BAB67794.1; ALT\_INIT.  
DR EMBL; AF155113; AAD42879.1; --  
DR Genew; HGNC:7744; NEK1.  
DR GK; Q96PY6; --  
DR MIM; 604588; --  
DR InterPro; IPR000719; prot kinase.  
DR InterPro; IPR008271; Ser Thr pkin.AS.  
DR InterPro; IPR002290; Ser Thr pkinase.  
DR InterPro; IPR001245; Tyr pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Prot kinase; 1.  
DR SMART; SM00220; S TKC; 1.  
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.  
DR PROSITE; PS0011; PROTEIN KINASE\_DOM; 1.  
DR TRANSFERASE; Serine/threonine-protein kinase; ATP-binding; Mitosis;  
KW Nuclear protein; Phosphorylation; Cell cycle; Cell division;  
KW Tyrosine-protein kinase; Alternative splicing.  
FT DOMAIN 4 258 PROTEIN KINASE.  
FT NP\_BIND 10 18 ATP (BY SIMILARITY).  
FT BINDING 33 33 ATP (BY SIMILARITY).  
FT ACT\_SITE 128 128 BY SIMILARITY.  
FT MOD\_RES 162 162 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
FT VARSPPLIC 478 521 Missing (in isoform 2).  
FT FTID=VSP 004870.  
FT CONFLICT 1232 1232 G -> E (IN REF. 2).  
FT SEQUENCE 1258 AA; 142828 MW; 339C4BF456612530 CRC64;

Query Match 23.3%; Score 582.5; DB 1; Length 1258;  
Best Local Similarity 29.9%; Pred. No. 7.6e-23;  
Matches 122; Conservative 93; Mismatches 124; Indels 69; Gaps 8;

QY 28 RYVLOKLGSGFTYLVSDKAKRGELKVLKETSVCGLNPNETVQANLAQLLSKLD 87  
DB 3 KYVRLQKIEGSGFKALLV--KSTEDGRQY-VIKELINISMKSKEERREVRVAVLANWK 59  
QY 88 HPAIVKFAFVQDNFCITTEYCEGRDLDKIQEYKQAGKIFPNQIIEWFIQLLLGVD 147  
DB 60 HPNIVQYRESFEENGSLYIVMDYCEGDLFKRIN--AQKGVLFQEDQILDWFVQICLALK 117  
QY 148 YMERILHDLKSKVEL-KNNLLKIGDGVSRLLMGSCDLATTLTGTPHYMSPEALKH 206  
DB 118 HVHDKILHDKISQNIFTKQGTQVQDGFGLARVLNLSVELARTICIGIPYLSPEICEN 177  
QY 207 QGYDTKSDIWSLACILYEMCMHAFAGSNFLSIVLKIYVEGDTPLSPERYKELNAIMES 266  
DB 178 KPTNKSDDIHALGCVLYELCTLKHAEAGSMKVLKLIISGFPVPSVLSHYSDLRSLVSQ 237  
QY 267 MLNKNLSRPSATEILKIPYLDQLQNLN-----C----- 296  
DB 238 LFKRNPDRPSVNSILEKGFIAKRIKFLSPQIAEEFLKTSFKGSPQIPAKRPASQ 297  
QY 297 -----RYSEMTEDKNDLQCKQAAHINAMQRIHLQTLR 331  
DB 298 NSISVMPAQKITPAKYGIPLAYKYKDGKGLHEKKPLQKHQAH--QFPEKRVN----- 350  
QY 332 ALSEVQMTPTREMRUKLQAADEKARKLKKIVVEKYENSKMQBLR 379  
DB 351 -TGEERRKISEEARKERLEFIEKKKQKQDQIISLMKASQMKRQEKER 397

RESULT 7  
ID NEK2\_HUMAN STANDARD; PRT; 445 AA.  
AC P51955;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Serine/threonine-protein kinase Nek2 (EC 2.7.1.37) (NIMA-related  
DE protein kinase 2) (NIMA-like protein kinase 1) (HSPK 21).  
GN NEK2 OR NLK1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Nasopharynx, Placenta, and T-cell;  
RX MEDLINE=94368699; PubMed=7522034;  
RA Schultz S.J., Fry A.M., Suetterlin C., Ried T., Nigg E.A.;  
RT "Cell cycle-dependent expression of Nek2, a novel human protein  
RT kinase related to the NIMA mitotic regulator of Aspergillus  
RT nidulans.";  
RL Cell Growth Differ. 5:625-635(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC Lu K.P., Hunter T.;  
RT "Molecular cloning and expression of NLK1, a human NIMA-like kinase.";  
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Skin;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Stapleton M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Brownstein M.J., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Raba S.S., Loquellano N.A., Malek J.A., Gunaratne P.H.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Heiton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).  
RN [4]  
RP SEQUENCE OF 83-203 FROM N.A.  
RX MEDLINE=94100173; PubMed=8274451;  
RA Schultz S.J., Nigg E.A.;  
RT "Identification of 21 novel human protein kinases, including 3 members  
RT of a family related to the cell cycle regulator nima of Aspergillus  
RT nidulans.";  
RL Cell Growth Differ. 4:821-830(1993).  
CC 1- FUNCTION: PROTEIN KINASE THAT IS INVOLVED IN MITOTIC REGULATION.  
CC MAY HAVE A ROLE AT THE G2-M TRANSITION. MAY ALSO PLAY A ROLE IN  
CC MEIOSIS.  
CC 1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.  
CC 1- SUBUNIT: Interacts with TERF1.  
CC 1- SUBCELLULAR LOCATION: Nuclear (By similarity).  
CC 1- DEVELOPMENTAL STAGE: ACCUMULATES THROUGHOUT S PHASE AND SHOWS  
CC MAXIMAL LEVELS IN LATE G2. THIS EXPRESSION PATTERN IS HIGHLY  
CC REMINISCENT OF THAT OF A AND B CYCLINS.  
CC 1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIMA  
CC subfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

```
CC DR EMBL; Z29066; CAA82309.1; -
DR EMBL; U11050; AAA19558.1; -
DR EMBL; BC043502; AAH43502.2; -
DR EMBL; Z25425; CAA80912.1; -
DR PIR; G01452; G01452.
DR PIR; I38215; I38215.
DR HSSP; Q00534; IBI7.
DR Genew; HGNC:7745; NEK2.
DR GIK; P51955; -.
DR MIM; 604043; -.
DR GO; GO:0005813; C:centrosome; TAS.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; TAS.
DR GO; GO:0007088; P:regulation of mitosis; TAS.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR TRANSFERASE; Serine/threonine-protein kinase; ATP-binding; Mitosis;
KW Nuclear protein; Phosphorylation; Cell cycle; Cell division; Meiosis.
FT DOMAIN 8 271 PROTEIN_KINASE.
FT NP_BIND 14 22 ATP (BY SIMILARITY).
FT BINDING 37 37 ATP (BY SIMILARITY).
FT ACT_SITE 141 141 BY SIMILARITY.
FT MOD_RES 175 175 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CONFLICT 84 85 IV -> LY (IN REF. 4).
SQ SEQUENCE 445 AA; 51763 MW; D33A37778ABB6D9E CRC64;

Query Match
Best Local Similarity 19.8%; Score 495; DB 1; Length 445;
Matches 139; Conservative 70; Mismatches 168; Indels 48; Gaps 14;

26 ARYVLQKLGSGSFGTVYLVSDKKAKGEEELKVL--KEISVGEINPNPNETVQANLEAQL 83
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 5 AEDYEVLVITGTSGYGRG-----QKIRKSDGKILVWKELDYGSMTEAKQMLVSEVNL 59
QY 84 SKLDHPAIVKHFASVEQDN--FCILTEYCEGRDLDDKIOEQKAGKIPPEINQIENFIQ 141
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 60 RELKHPNIVRYDRIIDRTNTLYIVMEYCEGGDLASVITKTKERQYLDDEFVLVVMVQ 119
QY 142 LLLGVVDYMER-----RIHLRLKSNVFL--KNLLKIGDFGVSRLMGSCDLATTLTG 194
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 120 LTLALKECHRRSDGGHTVLRDLKANVFLDGKQN-VKLGDFGLARILNHDTSFAKTVFG 178
QY 195 TPHYMSPEALKHQGYDTKSDIWSLACILYEMCOMHAFAGSNFLSIVLIVEGDTPLPE 254
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 179 TPYMSPEQMRYMSYNEKSDIWSLGLLYELCALMPPTAFSQELAGKIREGKFRIPY 238
QY 255 RYPKEINAIMESWLNKPSLRPSAIEILKIPVL-----DEQLQNLWCYSEMTLEDKNLD 309
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 239 RYSDNELIITMLNLDKYHRSVEIIEINPLADLVADEQRNLRERGRQLGEPKSD 298
QY 310 COKEAHINAVQ--KEIHLQ-----TIRALSE-----VQKMTPRMRRLKQLQADEKARK 359
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 299 -----SSPVLSELKKEIQLOERERALKAREERLEQKEQLCVRRLEADKLARENLLKN 354
QY 360 LKKIIVEEYKENSKEMLERSENFOOLSVDVLHEKTHLKGWEEK-----EQPEGRSLCS 414
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 355 YSLKERKFLSLANPELL-----NLPSSVITKKVHFSG-ESKENIMRSENSESLQTSK 407
QY 415 PQDED 419
DQ : : :
Db 408 SKCKD 412

RESULT 8
NEK8_BRARE
ID NEK8_BRARE STANDARD; PRT; 697 AA.
```

```
AC Q90XC2;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serine/threonine-protein kinase NEK8 (EC 2.7.1.37) (NimaA-related protein kinase 8).
GN NEK8.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Cyprinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]_TaxID=7955;
RP SEQUENCE FROM N.A.
RX MEDLINE=22308894; PubMed=12421721;
RA Liu S., Lu W., Obara T., Kuida S., Lehoczy J., Dewar K., Drummond I.A., Beier D.R.;
RT "A defect in a novel Nek-family kinase causes cystic kidney disease in the mouse and in zebrafish.";
RL Development 129:5839-5846(2002).
CC -!- FUNCTION: Required for renal tubular integrity.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- SUBCELLULAR LOCATION: Predominantly cytoplasmic (By similarity).
CC -!- DISEASE: Defects in NEK8 are the cause of polycystic kidney disease.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIMA subfamily.
CC -!- SIMILARITY: Contains 5 RCO1 repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF407580; AAL09676.1; -.
DR ZFIN; ZDB-GENE-020509.1; nek8.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR000408; Reg_Chrr_condens.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00415; RCC1; 3.
DR PRINTS; PR00633; RCNUNSATIOM.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00625; RCC1_1; FALSE_NEG.
DR PROSITE; PS00626; RCC1_2; FALSE_NEG.
DR PROSITE; PS00012; RCC1_3; 4.
DR Transferase; Serine/threonine-protein kinase; ATP-binding; Mitosis; Phosphorylation; Repeat.
KW DOMAIN 4 263 PROTEIN_KINASE.
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT REPEAT 417 468 RCC1 1.
FT REPEAT 469 520 RCC1 2.
FT REPEAT 521 586 RCC1 3.
FT REPEAT 587 636 RCC1 4.
FT REPEAT 637 689 RCC1 5.
FT ACT_SITE 128 128 BY SIMILARITY.
FT BINDING 33 33 ATP (BY SIMILARITY).
FT MOD_RES 162 162 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 697 AA; 76541 MW; 875A51D1F3831AA7 CRC64;

Query Match
Best Local Similarity 19.0%; Score 474.5; DB 1; Length 697;
Matches 111; Conservative 63; Mismatches 107; Indels 31; Gaps 7;
```



FT	BINDING	33	ATP (BY SIMILARITY).
FT	MOD_RES	162	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SEQ	SEQUENCE	692 AA; 74806 MW; 9E09820DFB3D5CA1 CRC64;	
	Query Match	18.2%; Score 454; DB 1; Length 692;	
	Best local similarity	36.1%; Pred. No. 1.5e-16;	
	Matches	97; Conservative 61; Mismatches 103; Indels 8; Gaps 4;	
QY	28	RYVLOQKLGSGFVGYLVSDKKAKRGEELKVLKEISVGEELPNFNETVQANLEAQLLSKLD 87	
DB	3	KYERIVVGRGAFIGVHLC--LRKADQKLVIIKQIPVQMTEKEERQAQNECVQLKLN 59	
QY	88	HPAIVKHFASVFEQDNFCITTEYCEGRDLDDTKTQYKQAGKIFPENQIIWFQILLGVD 147	
DB	60	HNVEVTEYENFLEPKALMIAMEYAPGCTLAEFTQ--KRCNSLIEBETILHFVQVILLALH 117	
QY	148	YMHERRIILHRDLKSNVFLKN--LLKIGDFGVSRLLMGSCDLATTLTCTPRYMSPEALK 205	
DB	118	VHVTHLILHRDLKTQNILDKHRNVKIGDFGSKIL--SSKAYTVVGTGTCYIISPELCE 176	
QY	206	HOGYDTKSDIWSLACILYEMCMNHPAGSNFLSIVLKIYVEGDTPLSPERYKPELNAIME 265	
DB	177	GKPNYQKSDIWALGCVLYELASLKRAPEAAANPALVLIKMTSGTFAPISDRYSPELRQVLV 236	
QY	266	SWLNKNPSLRPSAIEILKIPVLDEQLQNL 294	
DB	237	LSLEPAQRPLPSHIMQAQPLCTRALLNL 265	
RESULT 11			
NEK8 MOUSE			
ID	NEK8 MOUSE	STANDARD;	PRT; 698 AA.
AC	Q91ZE4; Q9D695;		
DT	10-OCT-2003 (Rel. 42, Created)		
DT	10-OCT-2003 (Rel. 42, Last sequence update)		
DT	10-OCT-2003 (Rel. 42, Last annotation update)		
DE	Serine/threonine-protein kinase Nex8 (EC 2.7.1.37) (NimA-related		
DE	protein kinase 8).		
GN	NEK8		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
NCBI	TaxId=10090;		
[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 1), MUTAGENESIS OF LYS-33, AND VARIANT		
RP	ARJPKD VAL-448;		
RP	STRAIN-C57BL/6J;		
RX	MEDLINE=22308894; PubMed=12421721;		
RA	Liu S., Lu W., Obara T., Kuida S., Lehoczy J., Dewar K.,		
RA	Drummond I.A., Beier D.R.;		
RT	"A defect in a novel Nek-family kinase causes cystic kidney disease in		
RT	the mouse and in zebrafish."		
RL	Development 129:5839-5846(2002).		
RP	[2]		
RP	SEQUENCE FROM N.A. (ISOFORM 2).		
RC	STRAIN-C57BL/6J; TISSUE=Skin;		
RC	MEDLINE=21085660; PubMed=11217851;		
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,		
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,		
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,		
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,		
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,		
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,		
RA	Kuehl P., Lewis S., Matsuo Y., Nikaide I., Resole G., Quackenbush J.,		
RA	Schriml L.M., Szaubli P., Suzuki R., Tomita M., Wagner L., Washio T.,		
RA	Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,		
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,		
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,		
RA	Gustincich S., Hill D., Hofmann C., Hume D.A., Kamiya M., Lee N.H.,		
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,		
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,		
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,		
RA	Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,		







```
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; kinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc_1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_SF; 1.
KW Transferrase; Serine/threonine-protein kinase; ATP-binding; Mitosis;
KW Cell cycle; Cell division.
FT DOMAIN 4 281 PROTEIN KINASE.
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 151 151 BY SIMILARITY.
SQ SEQUENCE 722 AA; 82689 MW; 20C7E304DAD7C440 CRC64;

Query Match 17.9%; Score 448; DB 1; Length 722;
Best Local Similarity 26.8%; Pred. No. 3.3e-16;
Matches 121; Conservative 96; Mismatches 169; Indels 66; Gaps 14;

QY 28 RYVLOKLGSGPTVYVSDKAKGEBELKVLKEISVGLNPNQVQANLEAQLSKLD 87
DB 3 KYKILECIGHGSGRTYKV--QRLKDG-ALLAQKEIHFGNITROBKQYIADENVILNLK 59
QY 88 HPAIVKPHASFVBDNFCT--ITEYCEGRDLDDKIQEYKQAGKIPENQIIEWFTQLLIG 145
DB 60 HPNIVQCGEELNRSQVNLNWEYCGHGLANLQRYKEEKRTQEVLKFFQLLLA 119
QY 146 VDYMH-----ERRILRLDKSNVFL-KNNLKIQDGFQVSLMGCS 185
DB 120 LYRCHYGENAPACDSQWPREIHPKQSVLHRDIKPAIEFDENNSVYKLDGFLSKLLDNT 179
QY 186 CDLATTLTCTPHYSKALHQGYDFTKSDIWLACILYEMCCWNAFAGSNFLSVLKVIV 245
DB 180 RVTQSVGVGTPYNSPEIRSSPYSDVWALGVIFECIMLTHPFGRSGYLEQRNIC 239
QY 246 EGTPTSLPERYPKELNAIMESMLNKNPSLRPAIEILKIPYLDQQLNLMCRYSMTLED 305
DB 240 QGNLSCWDHYSDDVFLIRHCLVNSDLRPTTYQLLRSPILSDIRSKL--ESERVLEQ 297
QY 306 KNIDCKOKEAAHINAMQKRIHLOTLRALSEVQKMTFRMRRLKQAADEKARKLKIVE 365
DB 298 SDL-----LHKH-QHMLIQENDLQ--FREQLSARESELENVIAISRLAQREE 342
QY 366 EKYEENSKRMOELSRNFQOLSVDVLHKLKMGEEKEEPEGRSLSCSPQDEDERWQG 425
DB 343 ILRRELEKQLRDMDAR-YQR-----HMQTVVNSMQKMR-----VTSFVDHNSQPSS 388
QY 426 REE-----ESDEPTLENLPE---SQPTSPM 447
DB 389 TAEVFDCTTEASQSPLHHPKLGISKPLQTL 420

RESULT 14
NEK7_MOUSE STANDARD; PRT; 302 AA.
AC Q9ES74;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serine/threonine-protein kinase Nek7 (EC 2.7.1.37) (Nima-related
DE protein kinase 7).
GN NEK7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RX SEQUENCE FROM N.A.
RX MEDLINE=20422672; PubMed=10964517;
RA Kandil M., Feige E., Chen A., Kilfin G., Motro B.;
RA "Isolation and characterization of two evolutionarily conserved murine
RT
```

```
RT kinases (Nek6 and Nek7) related to the fungal mitotic regulator,
RT NIMA";
RL Genomics 68:187-196(2000).
RP [2]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; and NOD; TISSUE=Thymus, and Urinary bladder;
RX MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Furuno M., Kasukawa T., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Nikaio I., Osato N., Saito R., Nogi A., Schonbach C., Gojibori T.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogi A., Schonbach C., Gojibori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Redierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Walestedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayaashiaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Czech II;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Heiton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIMA
CC subfamily.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
```



---

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 11:13:40 ; Search time 41 Seconds

(without alignments)  
1130.837 Million cell updates/sec

Title: US-10-803-278-4

Perfect score: 2501

Sequence: 1 MLKQFAAKCVSGSSTAIITY.....LGHHGDCNLISLDEYWKNEK 482

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:\*\*

1: p1r1:\*\*

2: p1r2:\*\*

3: p1r3:\*\*

4: p1r4:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	615	24.6	841	1 J78885	serine/threonine-s
2	600.5	24.0	792	2 JC7122	protein kinase (EC
3	598	23.9	774	2 S25284	protein kinase nek
4	495	19.8	445	2 G01452	NIMA-like protein
5	463.5	18.5	435	2 S23580	probable protein k
6	448	17.9	722	2 T37970	probable G2-specif
7	440.5	17.6	699	2 A43734	probable protein k
8	439	17.6	338	2 JC7838	Nek6 protein kinas
9	434.5	17.4	941	2 T49136	protein kinase-lik
10	428.5	17.1	357	2 T29771	hypothetical prote
11	414	16.6	294	2 T21075	hypothetical prote
12	410.5	16.4	779	2 A57177	NIMA-like protein
13	409.5	16.4	431	2 T11854	protein kinase (EC
14	392.5	15.7	1558	2 T29253	hypothetical prote
15	388	15.5	446	2 A40896	Ca2+/calmodulin-de
16	387	15.5	1233	2 T14157	serine/threonine p
17	382.5	15.3	1062	2 S43367	protein kinase CDC
18	376	15.0	848	2 T47986	serine/threonine-p
19	375	15.0	1231	2 T19532	serine/threonine pr
20	372.5	14.9	682	2 A44493	serum-inducible ki
21	370	14.8	200	2 B96587	hypothetical prote
22	368.5	14.7	447	2 B40896	hypothetical prote
23	362.5	14.5	795	2 JC4234	Ca2+/calmodulin-de
24	362	14.5	651	2 A96591	gene fused protein
25	358	14.3	1206	2 T34021	NPK1-related prote
26	355.5	14.2	1895	2 T15681	protein kinase SK2
27	354	14.2	603	2 S34130	hypothetical prote
28	353.5	14.1	1097	2 F96538	serine/threonine-s
29	345	13.8	1233	2 T30989	hypothetical prote
					serine/threonine p

30 344 13.8 469 2 B84644  
31 344 13.8 608 2 G96575  
32 342.5 13.7 445 2 T50802  
33 342 13.7 705 2 A48144  
34 340 13.6 426 2 S71886  
35 340 13.6 461 2 T48222  
36 340 13.6 603 2 A54596  
37 339.5 13.6 836 2 B96716  
38 337 13.5 273 2 S11380  
39 337 13.5 603 2 A47545  
40 336.5 13.5 974 2 S15038  
41 335.5 13.4 421 2 T48202  
42 335.5 13.4 738 1 S51380  
43 335.5 13.4 897 2 S61137  
44 335 13.4 435 2 B84707  
45 333 13.3 394 2 JUC0229

#### ALIGNMENTS

##### RESULT 1

I78885

serine/threonine-specific protein kinase (EC 2.7.1.1-) STK2 - human

C:Species: Homo sapiens (man)

C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 29-Sep-1999

C:Accession: I78885

R:Levedakou, E.N.; He, M.; Baptist, E.W.; Craven, R.J.; Cance, W.G.; Welcsh, P.L.; Simmor

Oncogene 9, 1977-1988, 1394

A:Title: Two novel human serine/threonine kinases with homologies to the cell cycle regul

A:Reference number: I58396; MUID:94268838; PMID:8208544

A:Accession: I78885

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-841 <RES>

A:Cross-references: GB:L20321; NID:g348244; PIDN:AAA36658.1; PID:g348245

C:Genetics:

A:Gene: GDB:STK2

A:Cross-references: GDB:374125

A:Map position: 3p21.1-3p21.1

C:Superfamily: human serine/threonine-specific protein kinase STK2; protein kinase homolo

C:Keywords: phosphotransferase

F:4-261/Domain: protein kinase homology <KIN>

Query Match 24.6%; Score 615; DB 1; Length 841;  
Best Local Similarity 34.7%; Pred. No. 3.3e-17;  
Matches 157; Conservative 87; Mismatches 152; Indels 56; Gaps 17;  
QY 29 YVQQKLGSGSFGTVYLVSDKKARGBELKYLKEISYGLNPNETVOANLEAQLSLKLDH 88  
Db 6 YCYLRVVGKSGYGEVTLV---KRRDQKQYVKKLNLRNASSRRERRAAEQEQLLSQLKH 62  
QY 89 PAIVKFIASFVEQDNFC-IITEYCEGRDLDKIQEYKQAGKIPENQIIEWFIQLLQVD 147  
Db 63 PNIVTYKESMEGGDGLLYVNGFCEGGDLYRKLKE--QKGQLLPENQVWFVQIAVALQ 120  
QY 148 YMHERRILHRDLKSNVFL-KNNLLKICDFGVSRLLMGSCDLATLTGTGTHYMSPEALKH 206  
Db 121 YLHEKHILHRDLKTVNQLRTNIIVKGLIGIARVLENHCDMASTLIGTPYKSPFLFSN 180  
QY 207 QGYDTKSDISLACILYEMCMNHAFAGSNFLSVLKIIVGDDTPSLPERTPKELNAMES 266  
Db 181 KPNYKSDVWALGCVVEMATLKAFNAKDMNSIVYRIIEGKLPAMPDRDYSPELAELIRT 240  
QY 267 MLNKNPILRPSAIEILKIPYLDEQLNLMCRYSEMTLED--KNLDCQ-KEAAHIN--- 319  
Db 241 MLSKEPERSVRSILRQPIKQISFLEATKIKTSKNNIKNGDSOSKPFATVVSGEAE 300  
QY 320 AMQKRIHQLTRA-----LSEYQKMTTPRMRRLKL--QAADKARKLKKIVVEKYEN 371  
Db 301 SNHEVTHPQLSSSGSQTYIMGEGKCLSQKPRASGLLKPSALKAKHTCQDL-----SN 355  
QY 372 SKNQELSRNFQQLSDVVLHEKTHLKGME-----KEEQPE-----GRLSCSPQD 417

Db 356 TTELATISSVN-----IDIL-----PAKGRDSVSDGFGVQENQPRYLDASNELGGI-CSISQ 405  
 QY 418 EDERWQGRREEDEPTLENLPESQIPSMDL 449  
 Db 406 VEEMLDQNTKSSAQF--ENL-----IPWSSDI 432

RESULT 2  
 JC7122  
 protein kinase (EC 2.7.1.37) 2 - mouse (strain balb/c)  
 N;Alternate names: serine (threonine) protein kinase  
 C;Species: Mus musculus (house mouse)  
 C;Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 31-Mar-2001  
 R;Hayashi, K.; Igarashi, H.; Ogawa, M.; Sakaguchi, N.  
 Biochem. Biophys. Res. Commun. 264, 449-456, 1999  
 A;Title: Activity and substrate specificity of the murine STK2 serine/threonine kinase  
 A;Reference number: JC7122; MUID:20001940; PMID:10529384  
 A;Accession: JC7122  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-792 <HAY>  
 A;Cross-references: GB:AJ223071; NID:G4138208; PID:G4138209  
 C;Genetics:  
 A;Gene: MSTK2L  
 C;Superfamily: human serine/threonine-specific protein kinase STK2; protein kinase homolog  
 C;Keywords: phosphotransferase

Query Match 24.0%; Score 600.5; DB 2; Length 792;  
 Best Local Similarity 31.5%; Pred. No. 1.2e-16;  
 Matches 141; Conservative 90; Mismatches 143; Indels 73; Gaps 13;

QY 29 YVLOQKLGSGFGTVLVSDKKAKRGEELKVLKEISVGLNPNETVQANLEAQLSKLDH 88  
 Db 6 YCMRVVGRSGSYGEVLV---KRRDQGVYIKKLNLRNASSRERRAARAEQAQLLSQLKH 62

QY 89 PAIVKPHASVFEODNFC--IITEYCEGRDLDDKIQEYKQAGKIPPENQIIEWFTQLLGYD 147  
 Db 63 PNIVTYKESWEGDGLLYIWMGCEGDLVYRKLKE--QKQLLPESQVVEFWQJAWALQ 120

QY 148 YMERRILHDLKSKNVFL--KNLLKIGDGVSRLLMGSCDLATTLTGTTPHMSPEALKH 206  
 Db 121 YLHEKHILHDLATQNVFLRTNIIKVGDLGARVLNHDGMASTLIGTPYNSPELFSN 180

QY 207 QGYDTKSDIWSLACILYEMCMNHAFAGSNFLSILVILKEGDPSPSPERYPKELNAIMES 266  
 Db 181 KPNYKSDVWALGCGCVEMATLKHAFNAKDNLSLVRIIEGKLPPMPKYVSTELAEILRT 240

QY 267 MLNKNPSLRPSAIEILKIPYLDLQ-----QNLKRYSEMVL-----ED 305  
 Db 241 MLSRRPERSVRSILRPQYIKKHISLFLFLEATKAKTSKNVKNCSRAKFAVAVSRKEE 300

QY 306 KNLD-----CQKRAAHIN-----AMQRIHLQTLRALSEVQKMTPRERMLRLKLA 352  
 Db 301 SNTDVIHYQPRSSGSAHVWGDKLCSQKPDVIGPLRSPLSGHTGKQDM-----NN 355

QY 353 ADEKARKLKKI-----VEEKYEENSKRMQELSRNFQQLSVDDLHEKTHLKGMEKEKEQ 406  
 Db 356 TGSCATISRINIDILPAERDSDANAGVQF-----SOPQHVDAADEVDVQ 401

QY 407 PEGRLSCSPQDEBERWQGRSESDP 433  
 Db 402 -----CSISQF-KERLQGNTKSSDQ 421

RESULT 3  
 S25284  
 protein kinase nek1 (EC 2.7.1.1) - mouse  
 C;Species: Mus musculus (house mouse)  
 C;Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 24-Sep-1999  
 R;Letwin, K.; Mizzen, L.; Motro, B.; Ben-David, Y.; Bernstein, A.; Pawson, T.

EMBO J. 11, 3521-3531, 1992  
 A;Title: A mammalian dual specificity protein kinase, Nek1, is related to the NIMA cell cycle  
 A;Reference number: S25284; MUID:93010942; PMID:1382974  
 A;Accession: S25284  
 A;Status: not compared with conceptual translation  
 A;Molecule type: mRNA  
 A;Residues: 1-774 <LET>  
 A;Cross-references: GB:S45928; NID:G256854; PIDN:AAB23529.1; PID:G256855  
 C;Genetics:  
 A;Gene: nek1  
 C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
 C;Keywords: ATP; phosphotransferase  
 F;2-58/Domain: protein kinase homology <KIN>  
 F;10-18/Region: protein kinase ATP-binding motif

Query Match 23.9%; Score 598; DB 2; Length 774;  
 Best Local Similarity 28.5%; Pred. No. 1.4e-16;  
 Matches 141; Conservative 94; Mismatches 132; Indels 128; Gaps 13;

QY 28 RYVLOQKLGSGFGTVLVSDKKAKRGEELKVLKEISVGLNPNETVQANLEAQLSKLD 87  
 Db 3 KYVRLQKIGSGFGKAVLV--KSTEDGRHY-VIKEINISRVSDKERQESRREAVAVLANMK 59

QY 88 HPATVPHASVFEODNFC--IITEYCEGRDLDDKIQEYKQAGKIPPENQIIEWFTQLLGYD 147  
 Db 60 HPNIVQYKESFEENGSLYIVMDYCEGDLFGRIN--AQKALFQEDQILDWFOVQICLALK 117

QY 148 YMERRILHDLKSKNVFL--KNLLKIGDGVSRLLMGSCDLATTLTGTTPHMSPEALKH 206  
 Db 118 HVDRKILHDLKSKNVFL--KNLLKIGDGVSRLLMGSCDLATTLTGTTPHMSPEALKH 177

QY 207 QGYDTKSDIWSLACILYEMCMNHAFAGSNFLSILVILKEGDPSPSPERYPKELNAIMES 266  
 Db 178 KPNYKSDIWSLACILYEMCMNHAFAGSNFLSILVILKEGDPSPSPERYPKELNAIMES 237

QY 267 MLNKNPSLRPSAIEILKIPYLDLQNLN-----C----- 296  
 Db 238 LFKRNPDRPSVNSILKGFIAKRIEFLSPQIAEFCLTKLSKFGQPLPKRPSAQ 297

QY 297 -----RYSEMTLEDKNDLDCQKEAAHIINAMQRIHLQTLR 331  
 Db 298 GVSSFVPAQKITKFAKYGVPLTYKYGDKKLEKXPPKHKQAHQI----- 344

QY 332 ALSEVQKMTPRERMLRLKLAQADEKARKLKIIVEEKYEENSKRMQELSRNFQQLSVDDL 391  
 Db 345 ---PVKXNNSGEERKKMSEAAA--KGRLEFIEKEK-----KQDKQIRFLKASQMK--R 391

QY 392 HEKTHLKGMEKEKEQ-----EGRLS---CSPQ----- 416  
 Db 392 QEKQLERINRAREQGWNVLRAGSGEVKASFPFGIGGAVSPSPSPRGQVHYHAFDQ 451

QY 417 -----DEBERWQ 425  
 Db 452 MQLRAEDNEARWKG 466

RESULT 4  
 G01452  
 NIMA-like protein kinase 1 - human  
 C;Species: Homo sapiens (man)  
 C;Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 24-Sep-1999  
 C;Accession: G01452  
 R;Lin, K.P.  
 Submitted to the EMBL Data Library, June 1994  
 A;Reference number: G07172  
 A;Accession: G01452  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-445 <LUX>  
 A;Cross-references: EMBL:U11050; NID:G507874; PIDN:AAA19558.1; PID:G507875  
 C;Genetics:  
 A;Gene: NIK1  
 C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog

C;Keywords: ATP

F;6-271/Domain: protein kinase homology <KIN>

F;14-22/Region: protein kinase ATP-binding motif

Query Match 19.8%; Score 495; DB 2; Length 445;  
Best Local Similarity 32.7%; Pred. No. 8.2e-13;  
Matches 139; Conservative 70; Mismatches 168; Indels 48; Gaps 14;

QY 26 ARRYVLOQKLGSGFGTVYLVSDKKARGBELKVL--KEISVGLGNPNETVQANLEAQLL 83

Db 5 AEDYEVLYTIGTSGYGRG-----QKIRKSDGKILVWKELDYGSMTEAEKQMLVSEVNL 59

QY 84 SKLDHPALVKFHFASVFEQDN--FCIIITEYCEGRDLDDKIQEYKQAKKIPPENOIIIEWFIQ 141

Db 60 RELKHPNIVRYDRIDRTNTTLYVMYCEGGDLASVITKTKERQYLDDEEFLVRVMTQ 119

QY 142 LLLGVDMYHHR-----RIHLRLDKSKNVFL--KNLLKIGDGVGSRLLMGSCDLATTLTG 194

Db 120 LTLALKECHRSDDGHTVLRDLKPNVFLDGKQN--VKLGDFGLAILNHDTSFATFVG 178

QY 195 TPHYMGPEALKHGYDTKSDIWSLACILYEMCCMHAFAGSNFLSVLKIVEGDTSPSLPE 254

Db 179 TPTYMSPEQNMNRSYNEKSDIWSLGCILLYELCALMPPTAFSQKELAGKIREGKFRIPY 238

QY 255 RYPKELNAINMESMLNKNPSLRPSAIELTKIPYL-----DEQLQNLWCRIYSEMTLEDKNLD 309

Db 239 RYSDLENEITRMNLKDHVRPSVEILENPLADIADVEQRNLRERRGRQGEPEKSD 298

QY 310 CQKEAAHIINAMQ--KRIHQ-----TTPALSE-----VQKMTPRERMRLKLAADKARK 359

Db 299 ----SSPVLSELKLSIQLERERALKAREERLEQKEQELCVREERLAEDKLARAENLLKN 354

QY 360 LKKIVEEKEENSQKMQELRSNFQQLSDVVLHEKTHLKGMEKE-----EQPEGRLSGS 414

Db 355 YSLKERRKFLSLASNPPELL-----NLPSSVTKKKVHFGS--ESKENIMRSENSESQLTSK 407

QY 415 PQDED 419

Db 408 SKCKD 412

RESULT 5

S23580  
probable protein kinase KIN3 (EC 2.7.1.1) - Yeast (Saccharomyces cerevisiae)  
N;Alternate names: probable protein kinase FUN52; probable protein kinase NPK1; protein  
C;Species: Saccharomyces cerevisiae

C;Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 24-Sep-1999

C;Accession: S23580; S20123; S40906; S43451; S11185

R;Schweitzer, B.; Philippsen, P.

Mol. Gen. Genet. 234, 164-167, 1992

A;Title: NPK1, a nonessential protein kinase gene in Saccharomyces cerevisiae with simil

A;Reference number: S23580; PMID:92357012; PMID:1495480

A;Accession: S23580

A;Molecule type: DNA

A;Residues: 1-435 <SCH>

A;Cross-references: EMBL:M67445; PID:g298024; PID:CAA43042.1; PID:g298026

R;Davies, C.J.; Hutchison III, C.A.

submitted to the EMBL Data Library, January 1992

A;Description: A directed DNA sequencing strategy based upon Tn3 transposon mutagenesis;

A;Reference number: S20123

A;Accession: S20123

A;Molecule type: DNA

A;Residues: 1-435 <DAV>

A;Cross-references: EMBL:M67445

R;Clark, M.W.; Keng, T.; Storms, R.K.; Zhong, W.; Fortin, N.; Zeng, B.; Delaney, S.; Oue

submitted to the EMBL Data Library, November 1993

A;Description: Sequencing of chromosome I of Saccharomyces cerevisiae: Analysis of the 4

A;Reference number: S40891

A;Accession: S40906

A;Molecule type: DNA

A;Residues: 1-430 <CLA>

A;Cross-references: EMBL:L22015; MIPS:VAR018C

R;Clark, M.W.; Keng, T.; Storms, R.K.; Zhong, W.; Fortin, N.; Zeng, B.; Delaney, S.; Oue

Yeast 10, 535-541, 1994

A;Title: Sequencing of chromosome I of Saccharomyces cerevisiae: analysis of the 42 kbp s

A;Reference number: S43441; PMID:95028152; PMID:7941740

A;Accession: S43451

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-435 <CLW>

A;Cross-references: EMBL:L22015; NID:G1339990; PIDN:AAO4964.1; PID:g349756

A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993

R;Jones, D.G.L.; Rosamond, J.

Gene 90, 87-92, 1990

A;Title: Isolation of a novel protein kinase-encoding gene from yeast by oligodeoxyribon

A;Reference number: S11185; PMID:90337351; PMID:2199332

A;Accession: S11185

A;Status: not compared with conceptual translation

A;Molecule type: DNA

A;Residues: 1-97, 'L' 99-120, 'Q' 122-129, 'I' 131-234, 'SQIS', 239-357, 'R' 359-369 <JON>

A;Cross-references: EMBL:M55416

A;Experimental source: strain SB303

C;Genetics:

A;Gene: SGD:KIN3; NPK1; FUN52

A;Cross-references: SGD:S0000071; MIPS:YAR018C

A;Map position: 1R

C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo

C;Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase

F;23-343/Domain: protein kinase homology <KIN>

F;31-39/Region: protein kinase ATP-binding motif

Query Match 18.5%; Score 463.5; DB 2; Length 435;

Best Local Similarity 31.1%; Pred. No. 1.3e-11;

Matches 127; Conservative 63; Mismatches 121; Indels 97; Gaps 12;

QY 29 YVLOQKLGSGFGTVYLVSDKKARGBELKVLKISVGLGNPNETVQANLEAQLLSKLDH 88

Db 25 YQVLEEIGRSGFSGVRKVIHPTKK--LLVRKDIKYGHMNSKEROQLIAECSSILSQLKH 81

QY 89 PALVKEHA-SFEVQ--DNFCIIITEYCEGRDLDDKIQEYKQAKKIPPENOIIIEWFI--QLLL 144

Db 82 ENIVEFYNDDEQREVLYMYEYCSRGDLSQMTKHVQEHKYIPEK--IVWGILAOILT 139

QY 145 GVDYMH-----ERRILHRLDKSKNVFLKN----- 169

Db 140 ALYKCHYGVELPTLTITVDRMKPPVKGKNIIVHRLKPGNIFLYSDSDYNNINEQVDGHE 199

QY 170 -----LLKGDGFGVSRLLMGSCDLATTLTGTPHYMSPGA 203

Db 200 EVNSNYVDHRVNSGKRGSPMDYSQVVVYKLGDFGLAKSLSTSIQFATTYVGTPTYMSPEV 259

QY 204 LKHQGYDTKSDIWSLACILYEMCCMHAFAGSNFLSVLKIVEGDTPSLPERYPKELNAI 263

Db 260 LMDQYPLSDIWSLGCVIFEMCSLHPFQAKNVLEQTKIKGKCDTVPYYSRGLNAI 319

QY 264 MESMLNKNPSLRPSAIELTKIPYLDEQLQNLWCRIYSEMTL-----EDKNLDCQKEAAHII 318

Db 320 IHSMDVNLTRPSTFEL-----LDQIQTARKSLQERFERFKLLDYENELTNIE 370

QY 319 NAMOKRIHQLTRALSQVQKWTPEERVLKLAADKARKL---KKI 363

Db 371 KILEKQA-IYERELSQJKE-----QPTQAVEERAREVISGKKV 408

RESULT 6

T37970

probable G2-specific protein kinase (EC 2.7.1.1) - fission yeast (Schizosaccharomyces pom

C;Species: Schizosaccharomyces pombe

C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Jan-2000

C;Accession: T37970

R;Badcock, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.

submitted to the EMBL Data Library, September 1997

A;Reference number: Z21758

A;Accession: T37970

A;Status: preliminary; translated from GB/EMBL/DBD

A;Molecule type: DNA





Db 291 YPPLPGEHYSEKLELYSMICIDPDPHQR 319  
| | | | | : : : : :  
RESULT 9  
T49136  
protein kinase-like protein - Arabidopsis thaliana  
N:Alternate names: protein P26G5.150  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000  
A:Accession: T49136  
R:D'Angelo, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Rudd, S.; I  
submitted to the Protein Sequence Database, April 2000  
A:Reference number: Z25017  
A:Accession: T49136  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-941 <DAN>  
A:Cross-references: EMBL:AL353814; GSPDB:GN00061; ATSP:F26G5.150  
A:Experimental source: cultivar Columbia; BAC clone F26G5  
C:Genetics:  
A:Gene: ATSP:F26G5.150  
A:Map position: 3  
A:Introns: 33/2; 55/3; 77/3; 93/2; 108/3; 135/3; 150/1; 166/3; 194/1; 210/3; 218/3; 237/3  
Query Match 17.4%; Score 434.5; DB 2; Length 941;  
Best Local Similarity 27.1%; Pred. No. 3.3e-10;  
Matches 138; Conservative 78; Mismatches 172; Indels 121; Gaps 19;  
QY 28 RYVLOQKLGSGFGTVLVSDKKAKGKELKYLKESVGEINPNETVQANLEAQLLSKLD 87  
| | | | | : : : : :  
Db 7 QYELMEQIGRGAGAAILVHKAERK---KYVUKKILRLARQTERCRRSAHQMSLIARVO 63  
| | | | | : : : : :  
QY 88 HPAIVKHFASFBODNF-CIITCYCEGRDLDDKIQYKQAGKIPFENQIIEWFIQLLV 146  
| | | | | : : : : :  
Db 64 HPYIVFEKAWKGVKVCIVTCYCEGGDAELMK--KSNGVYFPEKLCCKFTQLLAV 121  
| | | | | : : : : :  
QY 147 DYMHERRILHRDLKSNVFL-KNNILKIGDFGVSRLMGSCDLATLTGTPHMSPEALK 205  
| | | | | : : : : :  
Db 122 EYLSHNVVLRDLKCNIFLTKDQVRLGDFGLAKTLKAD-DLTSSVVGTPNYMCPPELLA 180  
| | | | | : : : : :  
QY 206 HQGYDTKSDTWSLACILYEMCCMNAFAGSNFLSIVLK--IVEGDPFSLPERYPKELNAI 263  
| | | | | : : : : :  
Db 181 DIPYGFKSDIWSLGCIIYEMAAVRPAKAFDMAGLSKSKSTHCQGVKEEPRYSAKR---- 236  
| | | | | : : : : :  
QY 264 MESMLNKNPSLRP-----SAEIL-----KPYLDEQLQNLKCRYSMTLE 304  
| | | | | : : : : :  
Db 237 MASEILKHPYLOPYVQYRPTLSAASITPEKPLNSREGRRSMAESQNSNSSEKKNFYVS 296  
| | | | | : : : : :  
QY 305 DKNL-----DCOKEAAHIINAMQKRIHLQTLRA-----LSEVQ 337  
| | | | | : : : : :  
Db 297 DKNIRVVPNGNKVTTETSGFVDDDEDLIDHVQOQSAENG-NIQSVSATKPDGHGILKPVH 355  
| | | | | : : : : :  
QY 338 K-----MTPRMRRLKLAQADEKARKLKIVKEEYKNSKRMQELRSR-----N 382  
| | | | | : : : : :  
Db 356 SQQRDPVIOPRPKTIIRNI-----MWVLKEEKARENGSPMSRNSRPSVPTQKN 406  
| | | | | : : : : :  
QY 383 FQQLSV-----DVLHE-----KTHLKMEZ--KEEQE 408  
| | | | | : : : : :  
Db 407 VETPSKIPKLGDIASHSSKTNASTPIPPSKLASDSARTPGSFPFKHMFVIDSSPKLPKN 466  
| | | | | : : : : :  
QY 409 GRLSCSP--QDEDEERWQREESDEPTL 435  
| | | | | : : : : :  
Db 467 DRISPSFAAKHAEAEAMSVKRRQRTPTL 495  
| | | | | : : : : :  
RESULT 10  
T29771  
hypothetical protein ZC581.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 17-Mar-2000  
C:Accession: T29771  
R:Waterston, B.; Gattung, S.; Le, T.T.

submitted to the EMBL Data Library, May 1997  
A:Description: The sequence of C. elegans cosmid ZC581.  
A:Reference number: Z20682

A:Accession: T29771  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-357 <WAT>  
A:Cross-references: EMBL:AF003134; PIDN:AB54139.1; GSPDB:GN00019; CESP:ZC581.1  
A:Experimental source: strain Bristol N2; clone ZC581  
C:Genetics:  
A:Gene: CESP:ZC581.1  
A:Map position: 1  
A:Introns: 31/3; 81/1; 120/3; 186/1; 226/3; 260/1; 288/3  
C:Superfamily: Ca2+/calmodulin-dependent protein kinase I; protein kinase homology

Query Match 17.1%; Score 428.5; DB 2; Length 357;  
Best Local Similarity 29.3%; Pred. No. 2.5e-10;  
Matches 96; Conservative 76; Mismatches 125; Indels 31; Gaps 7;

QY 35 LQSGSGFGTVLVSDKKAKGKELKYLKESVGEINPNETVQANLEAQLLSKLDHPAIVKF 94  
| | | | | : : : : :  
Db 10 VGRGAFGVWLCRGKNDASHQKV-IKLLINHTGHTKEENSIQSEVNLLKKVQHPILLIGY 68  
| | | | | : : : : :  
QY 95 HASFVEQDNFCITTEYCEGRDL-----DDKIQEYKQAGKIPFENQIIEWFIQLL 143  
| | | | | : : : : :  
Db 69 IDSFINDNQLGIVMAYAEQGTLERLINDQRAIKDSNMREY-----PPEKTVLDYFTQIL 122  
| | | | | : : : : :  
QY 144 LGVDYMHERRILHRDLKSNVFL--KNNLLKIGDFGVSRLMGSCDLATLTGTPHYMS 201  
| | | | | : : : : :  
Db 123 IALNNHQRKIVHRDLKSNVFL--KNNLLKIGDFGVSRLMGSCDLATLTGTPHYMS 181  
| | | | | : : : : :  
QY 202 EALKHQGYDTKSDIWSLACILYEMCCMNAFAGSNFLSIVLKIVEGDPFSLPERYPKELN 261  
| | | | | : : : : :  
Db 182 EICESRPNQKSDMWSLGCVLVYELLQLERAFDGENLPAIVMKITRSKQNLGLDHSVNDVK 241  
| | | | | : : : : :  
QY 262 AIMESMLNKNPSLRPSAEIILKIPYLBOLQNLKCRYSMTLEKDLDCQKEAAHIINAM 321  
| | | | | : : : : :  
Db 242 MLVENLLKTHTKRDPDVSQLLSDFLVPYLISIHCDLGR--IEPPPTDKRKPSASLSRL 299  
| | | | | : : : : :  
QY 322 QKRIHLQTLRLALS-----EVOKWTP 341  
| | | | | : : : : :  
Db 300 RTYPTQSTLRPYSLSNNAPTHTLTQLTP 327  
| | | | | : : : : :

## RESULT 11

T21075

hypothetical protein F19H6.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 17-Mar-2000

C:Accession: T21075; T21124

R:McMurray, A.

submitted to the EMBL Data Library, August 1995

A:Reference number: Z19368

A:Accession: T21075

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-294 &lt;WIL&gt;

A:Cross-references: EMBL:Z50873; PIDN:CAA90762.1; GSPDB:GN00028; CESP:F19H6.1

A:Experimental source: clone F17E5

R:McMurray, A.

submitted to the EMBL Data Library, November 1995

A:Reference number: Z19378

A:Accession: T21124

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-294 &lt;WIZ&gt;

A:Cross-references: EMBL:Z68115; PIDN:CAA92169.1; GSPDB:GN00028; CESP:F19H6.1

A:Experimental source: clone F19H6

C:Genetics:

A:Gene: CESP:F19H6.1

A:Map position: X

A:Introns: 55/3; 85/2; 113/3; 150/1; 192/3; 266/2

C:Superfamily: kinase-related transforming protein; protein kinase homology

Query Match 16.6%; Score 414; DB 2; Length 294;  
Best Local Similarity 34.9%; Pred. No. 7.6e-10;  
Matches 88; Conservative 56; Mismatches 90; Indels 18; Gaps 7;

QY 21 PKTLIARRVYLOQKLGSGFTVYLVSDKKARGE-----ELKV-LKEISVGLNPNETVQ 75  
DB 15 PDKLSLELFIIEKIKGQSEVF-----RAQCTWVDLHVALKIQVFMVDQKARQ 66

QY 76 ANL-EAQLLSKLDHPAIVKPHASVFDQDNFCITTEYCEGRDLDKIQEYKQAGKIPPNQ 134  
DB 67 DCLKEIDLLKQLNHNVNRYIASFIDNNQNLIVLELAEGDMSRMKPKKGRLLPEKT 126

QY 135 IIEWFTOLLGVDMYHERRILHRDLKSKNVFLK-NLLKIGDFGVSRLMGSCDLATTIT 193  
DB 127 IWKYFVQLARALAHMSKRIHMDIKPANVFTGNGIVKLGDLGLGRFFSSKTTAAHSIV 186

QY 194 GTPHYSPEALKHQGYDTSKDIWSLACILYEMCMNHAFAGS--NPLSIVLKIVEGDTPS 251  
DB 187 GTPYKSPERIQESGYNFKSGLWSTGCLLYEMAALQSPFYGDKNLYSLCKIENCEYPP 246

QY 252 LP-BRIPKELNA 262  
DB 247 LPADIVSTQVSA 258

RESULT 12  
A57177  
NIMA-like protein kinase - Neurospora crassa  
C:Species: Neurospora crassa  
C:Date: 10-Nov-1995 #sequence\_revision 10-Nov-1995 #text\_change 11-Aug-2003  
C:Accession: A57177  
R:Pu, R.T.; Xu, G.; Wu, L.; Vierula, J.; O'Donnell, K.; Ye, X.S.; Osmani, S.A.  
J. Biol. Chem. 270, 18110-18116, 1995  
A:Title: Isolation of a functional homolog of the cell cycle-specific NIMA protein kinase  
A:Reference number: A57177; MUID:95355415; PMID:7629122  
A:Status: preliminary; not compared with conceptual translation  
A:Accession: A57177  
A:Molecule type: DNA  
A:Residues: 1-779 <PUA>  
A:Cross-references: GB:L42573; NID:G1040682; PIDN:AAA80145.1; PID:G1040683  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
C:Keywords: ATP  
F:5-290/Domain: protein kinase ATP-binding motif  
F:13-21/Region:

Query Match 16.4%; Score 410.5; DB 2; Length 779;  
Best Local Similarity 31.2%; Pred. No. 2.4e-09;  
Matches 121; Conservative 65; Mismatches 151; Indels 51; Gaps 12;

QY 28 RYVLQKLGSGFTVYLVSDKKARGEELKVL--KEISVGLNPNETVQANLEAQLLSK 85  
DB 6 KYELLEKIGHSGFII-----RKVRRKADGMILCKEISYLKMSQKEREQLHAEFSLST 60

QY 86 LDHPAIV-KPHASVE-QDNFCITTEYCEGRDLDKIQEYKQAGKIPPNQIIEWFIQLL 143  
DB 61 LRHPNIVGYRHLKATQDLHLYMEVCGDGLGRVIRNLKKNQYAESEFWSIFSQVL 120

QY 144 LGVDYME-----ERRILHRDLKSKNVFL-KNLLKIGDF 176  
DB 121 TALYRCHGVDPPEVGKTVLGLSTARPKPSPGGMTILHRDLKPNVFLGEDNSVKLGDF 180

QY 177 GVSRLMGSCDLATLTCTPHYMSPEALKHQGYDTSKDIWSLACILYEMCMNHAFAGSN 236  
DB 181 GLSK-VMOSHDPASTVGTPTPMGPEICAAEKYLSKDIWSLGLIIVELCARPPFNKT 239

QY 237 FLSIVLKIVEGDTPSLPRYPKELNAIMESMLNKNPSLRPSAIELKIPYLDQQLNMC 296  
DB 240 HYQLVQKIKEGKIAFLPSVIGELFATIKDCLRVNPPRPDPTALLNLP-----IVLNR 294

QY 297 RYSENTLEDKLDCKEAAHIINAVQKEI-HQLTLRALSEVQKMTPRRMLRLKLOADE 355  
DB 295 KEKEVVEFSRTLTKEE-----TLNKRIRELDKSLGALETEKSSIRAEIDASLRREWEV 348

QY 356 KAR-KLKKIVVEEKYEE-NSKRMQELRSR 381  
DB 349 KARLEIDRLVAQRIEISLQKQFEQVQAR 376

RESULT 13  
T11854  
protein kinase (EC 2.7.1.1) - Trypanosoma brucei  
C:Species: Trypanosoma brucei  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 11-May-2000  
C:Accession: T11854  
R:Gale, M.J.; Parsons, M.  
Mol. Biochem. Parasitol. 59, 111-122, 1993  
A:Title: Trypanosoma brucei gene family encoding protein kinases with catalytic domains  
A:Reference number: 217363; MUID:93295429; PMID:8515773  
A:Accession: T11854  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-431 <GAL>  
A:Cross-references: EMBL:L03778; NID:G162169; PID:G162170  
C:Genetics:  
A:Note: nrka  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
C:Keywords: ATP; phosphotransferase

Query Match 16.4%; Score 409.5; DB 2; Length 431;  
Best Local Similarity 31.5%; Pred. No. 1.6e-09;  
Matches 85; Conservative 72; Mismatches 104; Indels 9; Gaps 5;

QY 28 RYVLQKLGSGFTVYLVSDKKARGEELKVLKEISVGLNPNETVQANLEAQLLSKLD 87  
DB 19 KYLXKGVIGLSYGEAYVA---ESVEDGSLCAVAVMDLSXMSQDKRYAQSEIKCLANCN 75

QY 88 HPAIVKPHASVFDQDNFCITTEYCEGRDLDKIQEYKQAG--KIFPENQIIEWFIQLL 145  
DB 76 HPNIIRYIEDHEENDRLIIVMEFADSGNLDEQI-KLRGSGDARYFQBHEALFQLQCLA 134

QY 146 VDYMERILHRDLKSKNVFL-KNLLKIGDFGVSRLMGSCD--LATTITGTTHYMSPE 202  
DB 135 LDYHSHKQHLHRDIKSNVLTSTGLVKLGDGFGSHQYEDTVSGVASTFCGTTPYLAPE 194

QY 203 ALKHQGYDTSKDIWSLACILYEMCMNHAFAGSNFLSIVLKIVEGDTPSLPRYPKELNA 262  
DB 195 LWNKRYNKYADVWSLGVLLYEMGMKPFSAASNLKGLMSKVLAGTYAPLPDSFSSEFKR 254

QY 263 IMESMLNKNPSLRPSAIELKIPYLDQQLQ 292  
DB 255 VVDGILVADPNDRPSVREIFQIPYINKGLK 284

RESULT 14  
T29253  
hypothetical protein B0496.3 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T29253  
R:Murray, J.; Lee, T.T.  
submitted to the EMBL Data Library, May 1996  
A:Description: The sequence of C. elegans cosmid B0496.  
A:Reference number: Z20596  
A:Accession: T29253  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1558 <MUR>  
A:Cross-references: EMBL:U58749; PIDN:AAB00636.1; GSPDB:GN00022; CESP:B0496.3  
A:Experimental source: strain Bristol N2; clone B0496  
C:Genetics:  
A:Gene: CESP:B0496.3  
A:Map position: 4  
A:Introns: 55/1; 100/3; 141/3; 226/3; 298/1; 468/3; 502/2; 609/3; 656/3; 795/3; 843/2; 90

Query Match 15.7%; Score 392.5; DB 2; Length 1558;



**This Page Blank (uspio)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2004, 11:12:20 ; Search time 121 Seconds  
(without alignments)  
1256.858 Million cell updates/sec

Title: US-10-803-278-4

Perfect score: 2501

Sequence: 1 MLKQFAAKCVSGSTAISTY.....LGVHGCNLSLDEYWKNEK 482

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_25.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mnc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2501	100.0	482	4	Q8TBY1	Q8TBY1 homo sapien
2	2408	96.3	470	4	Q8NG65	Q8NG65 homo sapien
3	2408	96.3	645	4	Q8NG66	Q8NG66 homo sapien
4	2310	92.4	637	6	Q8WNU8	Q8WNU8 macaca fasc
5	1757.5	70.3	628	11	Q8C004	Q8C004 mus musculus
6	1752.5	70.1	438	11	Q8BL56	Q8BL56 mus musculus
7	1596.5	63.8	365	11	Q8BW62	Q8BW62 mus musculus
8	1480	59.2	463	4	Q8H5F4	Q8H5F4 homo sapien
9	1396	55.8	395	6	Q8WNT6	Q8WNT6 macaca fasc
10	593.5	23.7	849	13	Q7T299	Q7T299 brachydanio
11	589.5	23.6	424	11	Q8CD72	Q8CD72 mus musculus
12	586.5	23.5	614	11	Q8CCJ0	Q8CCJ0 mus musculus
13	586.5	23.5	627	11	Q7TSC3	Q7TSC3 mus musculus
14	585.5	23.4	489	4	Q8J023	Q8J023 homo sapien
15	577	23.1	302	11	Q8BSB6	Q8BSB6 mus musculus
16	575.5	23.0	509	11	Q99K72	Q99K72 mus musculus

17	558	22.3	336	11	Q8C6N6	Q8C6N6 mus musculus
18	546.5	21.9	416	10	Q9LTP35	Q9LTP35 arabidopsis
19	543.5	21.7	1123	5	Q86I06	Q86I06 dictyosteli
20	514	20.6	943	10	Q94CU5	Q94CU5 cryza sativ
21	511	20.4	606	10	Q9CAU7	Q9CAU7 arabidopsis
22	509	20.4	621	10	Q8SA64	Q8SA64 populus x c
23	504	20.2	555	5	Q9N9C3	Q9N9C3 leishmania
24	499.5	20.0	555	10	Q8RXT4	Q8RXT4 arabidopsis
25	495	20.0	568	10	Q8RX66	Q8RX66 arabidopsis
26	495	19.8	491	4	Q86XH2	Q86XH2 homo sapien
27	495	19.8	609	10	Q947T1	Q947T1 lycopersico
28	485	19.4	442	13	Q9W622	Q9W622 xenopus lae
29	485	19.4	442	13	Q7ZYF3	Q7ZYF3 xenopus lae
30	476	19.0	384	4	Q9CQN9	Q9CQN9 homo sapien
31	475.5	19.0	841	5	Q9VC32	Q9VC32 dirosophila
32	472.5	18.9	440	13	Q7ZUN2	Q7ZUN2 brachydanio
33	468	18.7	389	13	Q9W623	Q9W623 xenopus lae
34	465.5	18.6	299	5	Q8T755	Q8T755 branchiosto
35	459	18.4	443	11	Q91Z18	Q91Z18 mus musculus
36	455.5	18.2	326	4	Q7Z634	Q7Z634 homo sapien
37	451	18.0	1057	5	Q815D5	Q815D5 plasmodium
38	450	18.0	443	11	Q921N9	Q921N9 mus musculus
39	450	18.0	579	5	Q95XQ3	Q95XQ3 caenorhabdi
40	445	17.8	310	13	Q7ZSX4	Q7ZSX4 xenopus lae
41	437.5	17.5	393	11	Q91XQ1	Q91XQ1 rattus norv
42	436.5	17.5	302	5	Q19530	Q19530 caenorhabdi
43	435	17.4	366	11	Q8BP64	Q8BP64 mus musculus
44	434.5	17.4	356	5	Q8N0P1	Q8N0P1 paramecium
45	434.5	17.4	941	10	Q9LXP3	Q9LXP3 arabidopsis

#### ALIGNMENTS

#### RESULT 1

Q8TBY1 ID Q8TBY1 PRELIMINARY; PRT; 482 AA.  
AC Q8TBY1;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
CX NCBI\_TaxID=9606;  
EN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RA Strausberg R.;  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EXBL: BC028587; AAH28587.1; -.  
DR GO: GO:000524; F:ATP binding; IEA.  
DR GO: GO:0004674; F:protein serine/threonine kinase activity; IEA.  
DR GO: GO:0004713; F:protein-tyrosine kinase activity; IEA.  
DR GO: GO:0016740; F:transferase activity; IEA.  
DR GO: GO:0006458; P:protein amino acid phosphorylation; IEA.  
DR InterPro: IPR000713; Prot kinase.  
DR InterPro: IPR002290; Ser\_Thr\_kinase.  
DR InterPro: IPR006271; Ser\_Thr\_pkin\_AS.  
DR InterPro: IPR001245; Tyr\_kinase.  
DR Pfam: PF00069; pkinase; 1.  
DR PRINTS: PR00109; TYRKINASE.  
DR ProDom: PD000001; Prot kinase; 1.  
DR SMART: SM00220; S\_TKc\_1.  
DR PROSITE: PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
DR Hypothetical protein; ATP-binding; Kinase;  
KW Serine/threonine-protein kinase; Transferase.  
SQ SEQUENCE 482 AA; 55519 MW; FA6D89B50C0223A5 CRC64;

Query Match 100.0%; Score 2501; DB 4; Length 482;

Best Local Similarity 100.0%; Pred. No. 2.3e-160;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLKFOBAKCVSGSTAIPTKTIARRVYVLOQKLGSGSGTGYLVSDKAKRGEELKVL 60  
DB 1 MLKFOBAKCVSGSTAIPTKTIARRVYVLOQKLGSGSGTGYLVSDKAKRGEELKVL 60

QY 61 KEISVGLNPNETVQANLEAQLSKLDHPAIVKHFASFEVDNFCIIITEYCEGRDLDDKI 120  
DB 61 KEISVGLNPNETVQANLEAQLSKLDHPAIVKHFASFEVDNFCIIITEYCEGRDLDDKI 120

QY 121 QYKQAGKIPENQIIIEWFIQLLGVDMYHERRILHRDLKSNVFLKNLLKIGDFGVS 180  
DB 121 QYKQAGKIPENQIIIEWFIQLLGVDMYHERRILHRDLKSNVFLKNLLKIGDFGVS 180

QY 181 LLMGSCDLATLTGTPTPHYMSPEALKHQGYDTKSDIWSLACILYEMCMNHAFAGSNFLSI 240  
DB 181 LLMGSCDLATLTGTPTPHYMSPEALKHQGYDTKSDIWSLACILYEMCMNHAFAGSNFLSI 240

QY 241 VLKIVEGDTPSLPERYPKELNAMESMLNKNPSLRPSAIEILKIPYLDOLQNLKCRYS 300  
DB 241 VLKIVEGDTPSLPERYPKELNAMESMLNKNPSLRPSAIEILKIPYLDOLQNLKCRYS 300

QY 301 MTLEDKNDLCCQKEAAHIINAMQKRIHLQTLRALSEVQKMTPRERMLRLKQAADERKAR 360  
DB 301 MTLEDKNDLCCQKEAAHIINAMQKRIHLQTLRALSEVQKMTPRERMLRLKQAADERKAR 360

QY 361 KKIIVEEYKNSKMOELSRNFQSLVDVHLKTHLKGMEKEBQEPGRLSCSPODEDE 420  
DB 361 KKIIVEEYKNSKMOELSRNFQSLVDVHLKTHLKGMEKEBQEPGRLSCSPODEDE 420

QY 421 ERWQGREESDEPTLENLPESQIPSPMDLHELESIVEDATSDLYGDCNLLSLDEYWN 480  
DB 421 ERWQGREESDEPTLENLPESQIPSPMDLHELESIVEDATSDLYGDCNLLSLDEYWN 480

QY 481 EK 482  
DB 481 EK 482

RESULT 2  
Q8NG65 PRELIMINARY; PRT; 470 AA.

ID Q8NG65  
AC Q8NG65  
DT 01-OCT-2002 (TRENBLrel. 22, Created)  
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE NIMA-related kinase 11S.  
GN NEK11S.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Noguchi K., Fukazawa H., Murakami Y., Uehara Y.;  
RT "Nek11, a new member of the NIMA family of kinases, involved in DNA  
RT replication and genotoxic stress responses."  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB071997; F:ATP binding; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
DR GO; GO:0016740; P:transferase activity; IEA.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
DR InterPro; IPR000719; Prot kinase.  
DR InterPro; IPR002290; Ser Thr kinase.  
DR InterPro; IPR008271; Ser Thr kinase.  
DR Pfam; PF00069; Pkinase; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Prot kinase; 1.  
DR SMART; SM00220; S\_TKc; 1.

DR SMART; SM00219; TyrKc; 1.  
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
KW ATP-binding; Kinase; transferase.  
SQ SEQUENCE 470 AA; 54006 MW; F52B82ED2B096FFB CRC64;

Query Match 96.3%; Score 2408; DB 4; Length 470;  
Best Local Similarity 100.0%; Pred. No. 4.1e-154;  
Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLKFOBAKCVSGSTAIPTKTIARRVYVLOQKLGSGSGTGYLVSDKAKRGEELKVL 60  
DB 1 MLKFOBAKCVSGSTAIPTKTIARRVYVLOQKLGSGSGTGYLVSDKAKRGEELKVL 60

QY 61 KEISVGLNPNETVQANLEAQLSKLDHPAIVKHFASFEVDNFCIIITEYCEGRDLDDKI 120  
DB 61 KEISVGLNPNETVQANLEAQLSKLDHPAIVKHFASFEVDNFCIIITEYCEGRDLDDKI 120

QY 121 QYKQAGKIPENQIIIEWFIQLLGVDMYHERRILHRDLKSNVFLKNLLKIGDFGVS 180  
DB 121 QYKQAGKIPENQIIIEWFIQLLGVDMYHERRILHRDLKSNVFLKNLLKIGDFGVS 180

QY 181 LLMGSCDLATLTGTPTPHYMSPEALKHQGYDTKSDIWSLACILYEMCMNHAFAGSNFLSI 240  
DB 181 LLMGSCDLATLTGTPTPHYMSPEALKHQGYDTKSDIWSLACILYEMCMNHAFAGSNFLSI 240

QY 241 VLKIVEGDTPSLPERYPKELNAMESMLNKNPSLRPSAIEILKIPYLDOLQNLKCRYS 300  
DB 241 VLKIVEGDTPSLPERYPKELNAMESMLNKNPSLRPSAIEILKIPYLDOLQNLKCRYS 300

QY 301 MTLEDKNDLCCQKEAAHIINAMQKRIHLQTLRALSEVQKMTPRERMLRLKQAADERKAR 360  
DB 301 MTLEDKNDLCCQKEAAHIINAMQKRIHLQTLRALSEVQKMTPRERMLRLKQAADERKAR 360

QY 361 KKIIVEEYKNSKMOELSRNFQSLVDVHLKTHLKGMEKEBQEPGRLSCSPODEDE 420  
DB 361 KKIIVEEYKNSKMOELSRNFQSLVDVHLKTHLKGMEKEBQEPGRLSCSPODEDE 420

QY 421 ERWQGREESDEPTLENLPESQIPSPMDLHELESIVEDATSDLYGDCNLLSLDEYWN 466  
DB 421 ERWQGREESDEPTLENLPESQIPSPMDLHELESIVEDATSDLYGDCNLLSLDEYWN 466

RESULT 3  
Q8NG66 PRELIMINARY; PRT; 645 AA.

ID Q8NG66  
AC Q8NG66  
DT 01-OCT-2002 (TRENBLrel. 22, Created)  
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE NIMA-related kinase 11L.  
GN NEK11L.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Noguchi K., Fukazawa H., Murakami Y., Uehara Y.;  
RT "Nek11, a new member of the NIMA family of kinases, involved in DNA  
RT replication and genotoxic stress responses."  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB071996; BAC06350.1; -.  
DR Genew; HGNC:18593; NEK11.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
DR GO; GO:0016740; P:transferase activity; IEA.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
DR InterPro; IPR000719; Prot kinase.  
DR InterPro; IPR002290; Ser Thr kinase.  
DR InterPro; IPR008271; Ser Thr kinase.  
DR ProDom; PD000001; Prot kinase; 1.  
DR SMART; SM00220; S\_TKc; 1.

DR Pfam, PF00069; pkinase; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Prot Kinase; 1.  
DR SMART; SM00220; S\_TK; 1.  
DR SMART; SM00219; TYK; 1.  
DR PROSITE; PS0011; PROTEIN KINASE DOM; 1.  
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.  
KW ATP-binding; Kinase; Transferase.  
SQ SEQUENCE 645 AA; 74162 MW; FE957805C5FFB059 CRC64;

Query Match 96.3%; Score 2408; DB 4; Length 645;  
Best Local Similarity 100.0%; Pred. NO. 6e-154;  
Matches 466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLKFOEAAKCVSGSTAISTYPTKTLIARRVYVLOQLGSGSGFTVYVSDKKARGBELKVL 60  
DB 1 MLKFOEAAKCVSGSTAISTYPTKTLIARRVYVLOQLGSGSGFTVYVSDKKARGBELKVL 60  
QY 61 KEISVGEINPNETVQANLEAQLSKLDHPAIVKPHASFEVDNFCIIITEYCEGRDLDDKI 120  
DB 61 KEISVGEINPNETVQANLEAQLSKLDHPAIVKPHASFEVDNFCIIITEYCEGRDLDDKI 120  
QY 121 QEYKQAGKIFPENQIIIEWFIQLLGVDMYHERRILHRDLKSNVFLKNNLLKIGDFGVSR 180  
DB 121 QEYKQAGKIFPENQIIIEWFIQLLGVDMYHERRILHRDLKSNVFLKNNLLKIGDFGVSR 180  
QY 181 LMGSCDLATLTGTTHYMSPEALKHQGYDTKSDIWSLACILYEMCCMNHAFAGSNFLSI 240  
DB 181 LMGSCDLATLTGTTHYMSPEALKHQGYDTKSDIWSLACILYEMCCMNHAFAGSNFLSI 240  
QY 241 VLKIVEGTPSLPERYPKELNAVESMLNKNPSLRPSAIEILKIPYLDLQNLNLCRYSE 300  
DB 241 VLKIVEGTPSLPERYPKELNAVESMLNKNPSLRPSAIEILKIPYLDLQNLNLCRYSE 300  
QY 301 MTEDEKNDLCKEAAHIINAMQKRIHLOTLRALSEVQKMTPRERMLRKLQAADKARKL 360  
DB 301 MTEDEKNDLCKEAAHIINAMQKRIHLOTLRALSEVQKMTPRERMLRKLQAADKARKL 360  
QY 361 KKIIVEEKEENSKMQLSRNFQOLSVDVLHEKTHLKGMEKEEQEGRGLSCSPQDEDE 420  
DB 361 KKIIVEEKEENSKMQLSRNFQOLSVDVLHEKTHLKGMEKEEQEGRGLSCSPQDEDE 420  
QY 421 ERWQGREESDEPTLENLPESQIPSMDLHELESIVEDATSDLYH 466  
DB 421 ERWQGREESDEPTLENLPESQIPSMDLHELESIVEDATSDLYH 466

RESULT 4  
Q8WN08 PRELIMINARY; PRT; 637 AA.  
ID Q8WN08  
AC Q8WN08;  
DT 01-MAR-2002 (TremBLrel. 20, Created)  
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)  
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)  
DE mRNA, similar to human hypothetical protein FLJ23495, complete cds.  
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecidae; Macaca.  
OX NCBI\_TaxID=9541;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RA Hashimoto K., Osada N., Hida M., Kusuda J., Tanuma R., Hirai M., Terao K., Suzuki Y., Sugano S.;  
RT "Isolation of novel full-length cDNA clones from macaque testis cDNA libraries."  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
KW EMBL; AB064997; BAB3539.1; -.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.

DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
DR GO; GO:0016740; F:transferase activity; IEA.  
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
DR InterPro; IPR000719; Prot\_Kinase.  
DR InterPro; IPR002290; Ser\_thr\_kinase.  
DR InterPro; IPR008271; Ser\_thr\_kinase.  
DR InterPro; IPR011245; Tyr\_kinase.  
DR Pfam; PF00069; pkinase; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR ProDom; PD000001; Prot\_kinase; 1.  
DR SMART; SM00220; S\_TK; 1.  
DR PROSITE; PS0011; PROTEIN KINASE DOM; 1.  
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.  
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
SQ SEQUENCE 637 AA; 73337 MW; FA3A5C3CEEAD16F CRC64;

Query Match 92.4%; Score 2310; DB 6; Length 637;  
Best Local Similarity 95.9%; Pred. NO. 2.3e-147;  
Matches 447; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 1 MLKFOEAAKCVSGSTAISTYPTKTLIARRVYVLOQLGSGSGFTVYVSDKKARGBELKVL 60  
DB 1 MLKFOEAAKCVSGSTAISTYPTKTLIARRVYVLOQLGSGSGFTVYVSDKKARGBELKVL 60  
QY 61 KEISVGEINPNETVQANLEAQLSKLDHPAIVKPHASFEVDNFCIIITEYCEGRDLDDKI 120  
DB 61 KEISVGEINPNETVQANLEAQLSKLDHPAIVKPHASFEVDNFCIIITEYCEGRDLDDKI 120  
QY 121 QEYKQAGKIFPENQIIIEWFIQLLGVDMYHERRILHRDLKSNVFLKNNLLKIGDFGVSR 180  
DB 121 QEYKQAGKIFPENQIIIEWFIQLLGVDMYHERRILHRDLKSNVFLKNNLLKIGDFGVSR 180  
QY 181 LMGSCDLATLTGTTHYMSPEALKHQGYDTKSDIWSLACILYEMCCMNHAFAGSNFLSI 240  
DB 181 LMGSCDLATLTGTTHYMSPEALKHQGYDTKSDIWSLACILYEMCCMNHAFAGSNFLSI 240  
QY 241 VLKIVEGTPSLPERYPKELNAVESMLNKNPSLRPSAIEILKIPYLDLQNLNLCRYSE 300  
DB 241 VLKIVEGTPSLPERYPKELNAVESMLNKNPSLRPSAIEILKIPYLDLQNLNLCRYSE 300  
QY 301 MTEDEKNDLCKEAAHIINAMQKRIHLOTLRALSEVQKMTPRERMLRKLQAADKARKL 360  
DB 301 MTEDEKNDLCKEAAHIINAMQKRIHLOTLRALSEVQKMTPRERMLRKLQAADKARKL 360  
QY 361 KKIIVEEKEENSKMQLSRNFQOLSVDVLHEKTHLKGMEKEEQEGRGLSCSPQDEDE 420  
DB 361 KKIIVEEKEENSKMQLSRNFQOLSVDVLHEKTHLKGMEKEEQEGRGLSCSPQDEDE 420  
QY 421 ERWQGREESDEPTLENLPESQIPSMDLHELESIVEDATSDLYH 466  
DB 421 ERWQGREESDEPTLENLPESQIPSMDLHELESIVEDATSDLYH 466

RESULT 5  
Q8C0Q4 PRELIMINARY; PRT; 628 AA.  
ID Q8C0Q4  
AC Q8C0Q4;  
DT 01-MAR-2003 (TremBLrel. 23, Created)  
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)  
DE Similar to mRNA.  
DN 4932416N14RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Testis;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium.  
RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of

60,770 full-length cDNAs.":  
 RT NATURE 420:563-573(2002).  
 RL EMBL; AK030042; BAC26756.1; -  
 DR MGD; MGI:2442276; 4932416N14RIK.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR000719; Prot kinase.  
 DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot kinase; 1.  
 DR SMART; SM00220; S\_TK; 1.  
 DR SMART; SM00219; TyK; 1.  
 DR PROSITE; PSS0011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 SQ SEQUENCE 628 AA; 71666 MW; AF6672BF80DCE71 CRC64;  
 Query Match 70.3%; Score 1757.5; DB 11; Length 628;  
 Best Local Similarity 73.4%; Pred. No. 3.7e-110;  
 Matches 343; Conservative 37; Mismatches 50; Indels 37; Gaps 2;  
 QY 1 MLKFOEAAKCVS-GSTAISTYPTKTLIARYVLOQKLGSGSGFTVYLVSDKKAKRGEELKV 59  
 DB 1 MLKFQETAKCVGRRTVPMYPTALIAIRYVLOQKLGSGSGFTVYLVSDKKAKRGEELKV 60  
 QY 60 LKEISVGLNPNETVQANLEAQLSLDHPAIVKPHASFEVDNFCIIITEYCEGRDLDDK 119  
 DB 61 LKEISVGLNPNETVQANVEAQLSLRHPAIVRPHASFEQETFCIIITEYCEGRDLDR 120  
 QY 120 IQEYKQAGKIPFENQIIEWFIQLLGVDMYHERRILHDLKSNVFLKNLLKIGDFGV 179  
 DB 121 IQEYKQAGKIPFENQIIEWFIQLLGVDMYHERRILHDLKSNVFLKNLLKIGDFGV 180  
 QY 180 RLIMGSCDLATLTGTPHYMSPEALKHOGYDTPKSDIWSLACILYEMCCMHAFAGSNFLS 239  
 DB 181 RLIMGSCDLATLTGTPHYMSPEALKHOGYDTPKSDIWSLACILYEMCCMHAFAGSNFLS 240  
 QY 240 IVLKIVGEGDTPSLPRYPKELNAIMESMLNKNPSLRPSAAILKAPYMEEQQLLMCKYP 299  
 DB 241 VLVNIVEGKTPSLPRYPRELNTIMERLNKSPSLRPSAAILKAPYMEEQQLLMCKYP 300  
 QY 300 EMTLEDKNLDCQKEAAHIINAMOKRIHLQTLRALSEVQKMTPRERMLRKLQAADKARK 359  
 DB 301 EMTLEDKNVCKQEAHTINAVQKLHLQTLQALSQTKTTPRERMLRKLQAADKARK 360  
 QY 360 LKKIVBEKYEENSKRMQELSRNFQOLSVDVLHETKHLKGMKEEKEPEGRSLSCSPQDED 419  
 DB 361 LKKIAEENYKENDKRMQALSRNVGSVHVLH----- 393  
 QY 420 EERWQGRSEESDEPTLENIPESQIPSMDLHELSIVEDATSLDGH 466  
 DB 394 -----ELDERTLESPEQSLPCLDLDELEPSLEDTIVDLGHY 431  
 RESULT 6  
 Q8BL56 PRELIMINARY; PRT; 438 AA.  
 ID Q8BL56  
 AC Q8BL56;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Similar to mRNA (Fragment).  
 GN 4932416N14RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RN SEQUENCE FROM N.A.

STRAIN-C57BL/6J; TISSUE-Cortex;  
 MEDLINE=22354683; PubMed=124666851;  
 RA THE FANTOM Consortium,  
 RT the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RA "Analysis of the mouse transcriptome based on functional annotation of  
 60,770 full-length cDNAs.":  
 RT NATURE 420:563-573(2002).  
 RL EMBL; AK03543; BAC31576.1; -  
 DR MGD; MGI:2442276; 4932416N14RIK.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.  
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.  
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro; IPR000719; Prot kinase.  
 DR InterPro; IPR002290; Ser\_thr\_pkinase.  
 DR InterPro; IPR008271; Ser\_thr\_pkin\_AS.  
 DR InterPro; IPR001245; Tyr\_pkinase.  
 DR Pfam; PF00069; pkinase; 1.  
 DR PRINTS; PR00109; TYRKINASE.  
 DR ProDom; PD000001; Prot kinase; 1.  
 DR SMART; SM00220; S\_TK; 1.  
 DR SMART; SM00219; TyK; 1.  
 DR PROSITE; PSS0011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 SQ SEQUENCE 438 AA; 50285 MW; DE540DB054F26D78 CRC64;  
 Query Match 70.1%; Score 1752.5; DB 11; Length 438;  
 Best Local Similarity 73.2%; Pred. No. 5.2e-110;  
 Matches 342; Conservative 37; Mismatches 51; Indels 37; Gaps 2;  
 QY 1 MLKFOEAAKCVS-GSTAISTYPTKTLIARYVLOQKLGSGSGFTVYLVSDKKAKRGEELKV 59  
 DB 1 MLKFQETAKCVGRRTVPMYPTALIAIRYVLOQKLGSGSGFTVYLVSDKKAKRGEELKV 60  
 QY 60 LKEISVGLNPNETVQANLEAQLSLDHPAIVKPHASFEVDNFCIIITEYCEGRDLDDK 119  
 DB 61 LKEISVGLNPNETVQANVEAQLSLRHPAIVRPHASFEQETFCIIITEYCEGRDLDR 120  
 QY 120 IQEYKQAGKIPFENQIIEWFIQLLGVDMYHERRILHDLKSNVFLKNLLKIGDFGV 179  
 DB 121 IQEYKQAGKIPFENQIIEWFIQLLGVDMYHERRILHDLKSNVFLKNLLKIGDFGV 180  
 QY 180 RLIMGSCDLATLTGTPHYMSPEALKHOGYDTPKSDIWSLACILYEMCCMHAFAGSNFLS 239  
 DB 181 RLIMGSCDLATLTGTPHYMSPEALKHOGYDTPKSDIWSLACILYEMCCMHAFAGSNFLS 240  
 QY 240 IVLKIVGEGDTPSLPRYPKELNAIMESMLNKNPSLRPSAAILKAPYMEEQQLLMCKYP 299  
 DB 241 VLVNIVEGKTPSLPRYPRELNTIMERLNKSPSLRPSAAILKAPYMEEQQLLMCKYP 300  
 QY 300 EMTLEDKNLDCQKEAAHIINAMOKRIHLQTLRALSEVQKMTPRERMLRKLQAADKARK 359  
 DB 301 EMTLEDKNVCKQEAHTINAVQKLHLQTLQALSQTKTTPRERMLRKLQAADKARK 360  
 QY 360 LKKIVBEKYEENSKRMQELSRNFQOLSVDVLHETKHLKGMKEEKEPEGRSLSCSPQDED 419  
 DB 361 LKKIAEENYKENDKRMQALSRNVGSVHVLH----- 393  
 QY 420 EERWQGRSEESDEPTLENIPESQIPSMDLHELSIVEDATSLDGH 466  
 DB 394 -----ELDERTLESPEQSLPCLDLDELEPSLEDTIVDLGHY 431  
 RESULT 7  
 Q8BW62 PRELIMINARY; PRT; 365 AA.  
 ID Q8BW62  
 AC Q8BW62;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Similar to mRNA (Fragment).  
 GN 4932416N14RIK.



```
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Oviduct;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK054237; BAC35699.1; -
DR MGD; MGI:2442276; 4932416N14rik.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0004648; F:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser Thr pkinase.
DR InterPro; IPR008271; Ser Thr pkin_AS.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot Kinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
FT NON_TER 365
SQ SEQUENCE 365 AA; 41752 MW; 01961635A759B7F2 CRC64;

Query Match 63.8%; Score 1596.5; DB 11; Length 365;
Best Local Similarity 83.8%; Pred. No. 1.3e-99;
Matches 305; Conservative 29; Mismatches 29; Indels 1; Gaps 1;

QY 1 MLKFOEAKCVS-GSTAISTYPTKTLIARRYVLQKLGSGSFGTVYLVSDKKAKGPEELKV 59
DB 1 MLKFOETAKCVGRRTPTVPMYPTALIARRYVLQKLGSGSFGTVYLVSDKKAKGPEELKV 60
QY 60 LKEISVGEINPNETVQANLEAQLSKLDHPATVKFHASVEODNFCITTEYCEGRDLDDK 119
DB 61 LKEISVGEINPNETVQANVEAQLSRHLHPALVRFHAFMEQETFCITTEYCEGRDLDYR 120
QY 120 IQEYQOAGKIIPPENQIIEFIIQLLGVDYMHERRILHRDLKSKNVLKIGDFGV 179
DB 121 IQEYKEAGVFAENQIVENFIQLLGVDYMHERRILHRDLKSKNVLKIGDFGV 180
QY 180 RLIMSGDLATLTGTTPHYMSPEALKHQGYDTKSDIWSLACILYEMCCMHAFAGSNFLS 239
DB 181 RLIMSGDLATLTGTTPHYMSPEALKHQGYDAKSDIWSLACILYEMCCLDHAFAGSNFLS 240
QY 240 IVLKIVEGDTPSLPERYPKELNAINESMLNKNPSLRPSAIEILKIPYLDEQLQNLNCRYS 299
DB 241 VVLNIVEGKTPSLPDYPRPELNTIMERMLNKNPSLRPSAAILKAPYMEQQLQNLNCRYP 300
QY 300 EMTLEDXNDQCKEAAHIINAMQRIHLOTLRALSEVQKMTPRERMLRKLQAADKARKK 359
DB 301 EMTLEDXNSVCKEAAHTINAVQKKLHLOTALSDTQKTPRERMLRKLQAADKARKK 360
QY 360 LKKI 363
DB 361 LKKI 364

RESULT 8
QH5F4 PRELIMINARY; PRT; 463 AA.
AC QH5F4
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
```

```
DE Hypothetical protein FLJ23495.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
RA Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T.,
RA Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y., Ota T., Suzuki Y.,
RA Obayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y.,
RA Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK027148; BAB15672.1; -
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004672; F:protein kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0004648; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot kinase; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Hypothetical protein; ATP-binding; Transferase.
SQ SEQUENCE 463 AA; 53403 MW; DEB585A8C1C939F9 CRC64;

Query Match 59.2%; Score 1480; DB 4; Length 463;
Best Local Similarity 100.0%; Pred. No. 1.2e-91;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 183 MGSCDLATLTGTTPHYMSPEALKHQGYDTKSDIWSLACILYEMCCMHAFAGSNFLSIVL 242
DB 1 MGSCDLATLTGTTPHYMSPEALKHQGYDTKSDIWSLACILYEMCCMHAFAGSNFLSIVL 60
QY 243 KIVEGDTPSLPERYPKELNAINESMLNKNPSLRPSAIEILKIPYLDEQLQNLNCRYSMT 302
DB 61 KIVEGDTPSLPERYPKELNAINESMLNKNPSLRPSAIEILKIPYLDEQLQNLNCRYSMT 120
QY 303 LEDKNLDCQKEAAHIINAMQRIHLOTLRALSEVQKMTPRERMLRKLQAADKARKLKK 362
DB 121 LEDKNLDCQKEAAHIINAMQRIHLOTLRALSEVQKMTPRERMLRKLQAADKARKLKK 180
QY 363 IVEKYEENSKRWQELSRNFQQLSVDLVHEKHLKGMEEKKEQPEGRILSCSPQDEDEER 422
DB 181 IVEKYEENSKRWQELSRNFQQLSVDLVHEKHLKGMEEKKEQPEGRILSCSPQDEDEER 240
QY 423 WQGREESDEPTLENLPESQIPISMDLHELESIVEDATSDLYH 466
DB 241 WQGREESDEPTLENLPESQIPISMDLHELESIVEDATSDLYH 284

RESULT 9
QH5NT6 PRELIMINARY; PRT; 395 AA.
ID QH5NT6
AC QH5NT6
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE mRNA, similar to human hypothetical protein FLJ23495, complete cds.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Hashimoto K., Osada N., Hida M., Kusuda J., Tanuma R., Hirai M.,
RA Terao K., Suzuki Y., Sugano S.;
RT "Isolation of novel full-length cDNA clones from macaque testis cDNA
libraries.";
```

RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB066559; BAB84034.1; -;  
 DR GO: GO:0005524; F-ATP binding; IEA.  
 DR GO: GO:0004672; F-protein kinase activity; IEA.  
 DR GO: GO:0016740; F-transferase activity; IEA.  
 DR GO: GO:0006468; P-protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR000719; Prot kinase.  
 DR Pfam: PF00069; pkinase; 1.  
 DR ProDom: PD000001; Prot kinase; 1.  
 DR PROSITE: PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR ATP-binding; transferase.  
 KW SEQUENCE 395 AA; 45391 MW; 585EF13B2B89B85 CRC64;  
 SQ SEQUENCE 395 AA; 45391 MW; 585EF13B2B89B85 CRC64;  
 Query Match 55.8%; Score 1396; DB 6; Length 395;  
 Best Local Similarity 94.7%; Pred. No. 4.6e-86;  
 Matches 269; Conservative 9; Mismatches 9; Indels 0; Gaps 0;  
 QY 183 MGSCDLATLTGTGPHNSPEALKHQGYDTSKDIWSLACILYEMCCNHAFAGSNFLSIVL 242  
 DB 1 MGSCDLATLTGTGPHNSPEALKHQGYDTSKDIWSLACILYEMCCNHAFAGSNFLSIVL 60  
 QY 243 KIVEGTPSLPERYPKELNAIMESMLNKNPSLRPSAIELKIPYLDLQNLNLCRYSEMT 302  
 DB 61 KIVEGTPSLPERYPKELNAIMESMLNKNPSLRPSAIELKIPYLDLQNLNLCRYSEMT 120  
 QY 303 LEDKNDCKEAAHINAMQRIHLQTLRALSEVQMTPREMRLKQAADKARKLKK 362  
 DB 121 LEDKNDCKEAAHINAMQRIHLQTLRALSEVQMTPREMRLKQAADKARKLKK 180  
 QY 363 IVEEKEEENKMOELRSNFQSLVDVLHKLGMKEEKEEPEGRSCSPQDEDEER 422  
 DB 181 IVEEKEEENKMOELRSNFQSLVDVLHKLGMKEEKEEPEGRSCSPQDEDEER 240  
 QY 423 WQREESDEPTLENIPESQIPSPMDLHELESIVEDATSDLYH 466  
 DB 241 WQREESDEPTLENIPESQIPSPMDLHELESIVEDATSDLYH 284  
 RESULT 10  
 QY7299 PRELIMINARY; PRT; 849 AA.  
 AC QY7299;  
 DT 01-OCT-2003 (T-EMBLrel. 25, Created)  
 DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Donald T.L., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Heiton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RP "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RA Strausberg R.;  
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC054633; AAH54633.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 849 AA; 95966 MW; E4EBB117A0147D7B CRC64;  
 Query Match 23.7%; Score 593.5; DB 13; Length 849;  
 Best Local Similarity 29.3%; Pred. No. 1.3e-31;  
 Matches 141; Conservative 92; Mismatches 178; Indels 71; Gaps 11;  
 QY 29 YVLOQKLGSGFVTVLVSDDKAKRGEELKVLKEISVGLNPNETVQANLEAQLSKLDH 88  
 DB 4 YLFVTVVVGKSGYGEVNLVRHKSDBK---QVVIKLNLRITSSRRERAAQEAQLLSQLKH 60  
 QY 89 PAIVKFHASVFOQ-NFCIITEYCEGRDLDKTIQYKQAKIPENQIIEWFIQLLQVD 147  
 DB 61 FNIWVRESWEGEDCQIYIVMGFCGGDLVHRLKQ--QKGEELPERQVVEWVQIAMAQ 118  
 QY 148 YMHEERTILHRDLKSNVFL-KNLLKIGDFGVSRLMGSCDLATLTGTGPHNSPEALKH 206  
 DB 119 YLHEKHILHRDLKTONIFLTKTNIIVKVDLGIAVLNQNNDMASTLIGTFYMSPLFSN 178  
 QY 207 QGYDTKSDIWSLACILYEMCCNHAFAGSNFLSIVLKIIVEGTPSLPERYPKELNAIMES 266  
 DB 179 KPYNKSDVWALGCVVEMATLKAFNAKDMNLVRIVEGKLPMPSKYDPOLGELIKR 238  
 QY 267 MLNKNPSLRSAIELKIPYLDLQNLNLCRYSEMT----- 302  
 DB 239 MLCKKPEDRPDVKHIDRQPIKHQISNLFLEATKAKRKNNAAGKLSAGSDATKPN 298  
 QY 303 -----LEDKNDCKEAAHINAMQK-----RIHLQTLRALSEVQMTPRE 344  
 DB 299 QVWVQPCLNSEKTCGKKAEEIYLNKQPCNGAWENVAPKHHMPKSPTRDIENSTGSI 358  
 QY 345 MRLKQAADKARKLKIIVEEKEEN-----SKRMOELRSNFQSLVDVLHKLHKG 399  
 DB 359 ATISNIDIEIQKQKAKPPKPPSHNNLPSVSKR-----REKDPGAPQTHPHKQVSG 412  
 QY 400 MEEKEEPEGRKL-SCSPQDEDEERWQGREESD-----EPTLENLPES---QIPSPMD 448  
 DB 413 VGGTEDIKMSANASSITPKPADRTKMPNKSALDVLSDIKDDTMKLQEAQVQDLPVEPTE 472  
 QY 449 LH 450  
 DB 473 LH 474  
 RESULT 11  
 QY8CD72 PRELIMINARY; PRT; 424 AA.  
 AC QY8CD72;  
 DT 01-MAR-2003 (T-EMBLrel. 23, Created)  
 DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)  
 DE NINA.  
 GN NEKI.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 60,770 full-length cDNAs.";

```

RL Nature 420:563-573 (2002).
DR EMBL; AK031330; BAC27350.1; -.
DR MGD; MGI:97303; Nekl.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TK; 1.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
SQ SEQUENCE 424 AA; 48635 MW; 416AD0E55C13CEFE CRC64;

Query Match 23.6%; Score 589.5; DB 11; Length 424;
Best Local Similarity 30.3%; Pred. No. 1e-31;
Matches 132; Conservative 88; Mismatches 128; Indels 87; Gaps 10;

QY 28 RYVLOQLGSGFQTVLYSDKAKRGELKVLKEISVSGELNPNETVQANLEAQLLSKLD 87
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 3 KYVRLQKIGEGFGKAVLY--KSTEDGRHY-VIKESINSRMSDKRQSSRREVAVLANMK 59
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 88 HPAIVKHFASFEVQONFCIITYCEGRDLDDKIQYKQAGKIFPENQIIEFIQLLQVD 147
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 60 HENIVQYKESFPENGSLYIMVDCGGDLFKEN--AQKALFQEDQILDWPFVQICLALK 117
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 148 YMHERILHRDLKSNVFL-KNNLLKIGDFGVSRLMGSCDLATLTGTTPHYMSPEALKH 206
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 118 HVHDKILHRDIKSNIFLTQDGTQGLGDFGIARVLNSTVELARTCIGTPYVLSPEICEN 177
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 207 QGYDVKSDIWSLACILYECMNHAFAGNSFIVLKIVEGDTFSLPERYPKELNAIMES 266
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 178 KPYNKSIDIALGCVLYELCTLKHAFAGNKNVLKLIISGFPVPVSHYSYDLRSLSIQ 237
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 267 MLNKNPSLRPSAIEILKIPYLDQLQNLN-----C----- 296
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 238 LFKRNPDRPVSNSILEKGFIAKRIEKLPSQLIABEFCCLKTSKFGPQLPGKRPASGQ 297
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 297 -----RYSEMTLEDKILDCQKEAAHIIINAMQKRIHLQILR 331
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 298 GYSSVFPAQKITPAKYGVPLTYKYGDKLEKPPPKHQAHOI----- 344
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 332 ALSEVQKMTPRMRRLKLADEKARKLKIVVEEYENSKRMQELSRNPFQQLSVDVL 391
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 345 ---PVKNNSGEEKWSEEA--KKRLEFIEKEK-----KQKQIRFLKAEQMK--R 391
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 392 HEKTHLKGMEKEEQ 406
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 392 QEKQRLERINRAEQ 406
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 12
Q8CCJ0 PRELIMINARY; PRT; 614 AA.
ID Q8CCJ0
AC Q8CCJ0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Weakly similar to protein kinase nekl.
GN BB049667.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
EA the RIKEN Genome Exploration Research Group Phase I & II Team;
ET "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573 (2002).";
DR EMBL; AK032672; BAC27980.1; -.
DR MGD; MGI:2142824; BB049667.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TK; 1.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
SQ SEQUENCE 614 AA; 70032 MW; DA95120EA45D71BC CRC64;

Query Match 23.5%; Score 586.5; DB 11; Length 614;
Best Local Similarity 31.8%; Pred. No. 2.6e-31;
Matches 148; Conservative 91; Mismatches 145; Indels 81; Gaps 16;

QY 35 LGSSGFTGVLYSDKAKRGELKVLKEISV--GELNPNETVQANLEAQLLSKLDHPAI 91
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 10 IGEFTKGVLYAKDKSE---SSHCIVKEISLTKESAKNEVI-----LLARMEHPNI 59
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 92 VKHFASFEVQDNFCIITYCEGRDLDDKIQYKQAGKIFPENQIIEFIQLLQVDVYME 151
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 60 VTFPSSFOENGRLFIVMYCDGGDLMOIQ--RQRGWVFSQDQLCFVQISLGLKHIHD 117
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 152 RRIILHRDLKSNVFLKNN--LLKIGDFGVSRLMGSCDLATLTGTTPHYMSPEALKHGY 209
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 118 RRIILHRDIKSNIFLPSKNGVAKLGDFTGARTLNDSELAQTACGTGTYVLSPEICQNPY 177
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 210 DTKSDIWSLACILYECMNHAFAGNSFIVLKIVEGDTFSLPERYPKELNAIMESMLN 269
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 178 NKTIDISLGLCVLYELCTLKHFPESNHNHFLVLKICQGRVAPISPHFSRDLQSLIPOLFR 237
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 270 KNPRLPSAIEILKIPYLDQLQNLNMC--YSEMTLEDKILDCQKEA--AHINA----- 320
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 238 VSPQDRPESVTSLLKRPF---LETLIARSLYPEV-----CSRITQSHARMENMAIGP 285
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 321 -----MQRHILQTLRALSEVQ-----KMTPRRMLR-----KLQAA 353
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 286 TACWVSFWSNAYLQRFKAQYKLVKVERQLGLRPPSVSEPHNEGEKQLQSHWEETKFQEL 345
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 354 DEKARKLK-----KIVVEEYENSKRMQELSRNPFQQLSVDVLHEKTHLKGMEF-----KE 404
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 346 QYRKNKMDQBYWKQLEIRQQYVHNDMKEIKKMGRELKRVKPEISLDCIASEDTVQE 405
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 405 EOPSGRLSCSPQDED-----EERWQGREESDEPTLENL--PESQ 442
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 406 NEAVDKLNATLSFEDGTGTFQFHRCKEHEHDYTDRAFEELCGPEAE 450
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 13
Q7TSC3 PRELIMINARY; PRT; 627 AA.
ID Q7TSC3
AC Q7TSC3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Olfactory epithelium;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hsieh L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ueding T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Richards S., Worley K.C., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RL cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RP [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Olfactory epithelium;  
 RX Strausberg R.;  
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC053516; AAH53516.1; --  
 KW Hypothetical protein  
 SQ SEQUENCE 627 AA; 71790 MW; AD5789B6621AC2P CRC64;

Query Match 23.5%; Score 586.5; DB 11; Length 627;  
 Best Local Similarity 31.8%; Pred. No. 2.7e-31;  
 Matches 148; Conservative 91; Mismatches 145; Indels 81; Gaps 16;

QY 35 LGSFGTGVYVSDKKAKEBELKVLKESV---GELNPETVQANLEAQLSLKLDHAI 91  
 DB 10 IGBGTGKYLAKDKSE---SHCVKIEISLTKYKAEAKNEVI-----LLARMEHNI 59  
 QY 92 VKFHAFVQDNFCITTEYCEGRDLDKIOEYKQAGKIPFENQIIEWFIQLLGVYDYMHE 151  
 DB 60 VTFSSFGNGRLFIWMEYCDGDLQRIQ---RQRGVMPSEQILCWVQISLGLKTHD 117  
 QY 152 RAILHRDLKSKVFLKN---LLKIGDGVSRLLMGSCDLATTLTGTPHYMSPEALKHQY 209  
 DB 118 RKLHRLDIKQNIIFSKNGWAKLGDGFGPARTLNDSMELAQTCAGTPYLSPEICQNPY 177  
 QY 210 DTKSDIWSLACILYEMCCMHAFAGSNFLSIVLKIIVEGDTPLSPRYPKELNAMESLN 269  
 DB 178 NKKTDIWSLGCVLIELCTLKHFPESNFHVLKICQGEVAPISPHSRDLQSLPQLFR 237  
 QY 270 KNPSPRPSAIEILKIPYDEQLMCR---YSEMTLEDKNDLQCKEA---AHINA----- 320  
 DB 238 VSPQDRPSVTSLLKRPF---LETIARSLYPEV-----CSRRIOSHAMMENAI GP 285  
 QY 321 -----MOKRIHLQTLRALSEVQ-----KMTPRMRRLR-----KLQAA 353  
 DB 286 TACWRVSPNSAAYLQKFAQYKLVKQRLGRSPSSVEPHNEGKELQSHWEETKFQEL 345  
 QY 354 DEKARKLK-----KIVVEKYENSKRMQELRNFTQQLSDVVLHKTGLKGNVE-----KE 404  
 DB 346 QYRKNNKMDQYWKQLEIRQYVHNDMEIKKMGRELKRVVVKFPEISLDKCISEEDTVQE 405  
 QY 405 EQPEGLSCSPQDE-----ERRWQGRREESDEPTLENI---PESO 442  
 DB 406 NEADVKNATLSFEDGTGKQHRCKEHEHEDYTDRAFEELCGPEAE 450

RESULT 14  
 Q8J023  
 ID Q8J023 PRELIMINARY; PRT; 489 AA.  
 AC Q8J023;  
 DT 01-MAR-2003 (TREMBlrel. 23, Created)  
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE NIMA-related protein kinase 3.  
 GN NEK3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kimura M., Okano Y.;  
 RL "Molecular cloning of human NEK3";  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB072828; BAC15599.1; --  
 DR GO: GO:0005524; F:ATP binding; IEA.  
 DR GO: GO:0004874; F:protein serine/threonine kinase activity; IEA.  
 DR GO: GO:0004713; F:protein-tyrosine kinase activity; IEA.  
 DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.  
 DR InterPro: IPR000719; Prot kinase.  
 DR InterPro: IPR002271; Ser Thr\_pkinase.  
 DR InterPro: IPR001245; Tyr\_pkinase.  
 DR Pfam: PF00069; pkinase;  
 DR PRINTS: PRO0109; TYRKINASE.  
 DR PRODOM: PD000001; Prot\_kinase; 1.  
 DR SMART: SM00220; S\_TKc; 1.  
 DR SMART: SM00219; TyKc; 1.  
 DR PROSITE: PS00107; PROTEIN KINASE ATP; 1.  
 DR PROSITE: PS00011; PROTEIN KINASE DOM; 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW Kinase.  
 SQ SEQUENCE 489 AA; 55983 MW; FAESDAAC01EFA88A CRC64;

Query Match 23.4%; Score 585.5; DB 4; Length 489;  
 Best Local Similarity 31.4%; Pred. No. 2.3e-31;  
 Matches 151; Conservative 87; Mismatches 154; Indels 89; Gaps 15;

QY 29 YVLOQKLGSGFTVILVSDKKAKEBELKVLKESVGEINPNETVQANLEAQLSLKLDH 88  
 DB 4 YWLVMTGEGFGRALLVQHSS---NQMPAMKEILPKFSFN-TQNSKEAVLLAKMKH 59  
 QY 89 PAIVKPHASVQDNFCITTEYCEGRDLDKIOEYKQAGKIPFENQIIEWFIQLLGVYD 148  
 DB 60 PNIVAFKSEFAGHLYVMEYCDGDLQRIQ---QKGLFPEDMLNWFQVCLGVNH 117  
 QY 149 MHERRILHRDLKSKVFL-KNLLKIGDGVSRLLMGSCDLATTLTGTPHYMSPEALKHQ 207  
 DB 118 IHKRVLRHDIKSKNIFLTQNGKVLGDFGSAARLLSNPMFACTVYVTPYVPEIWEHL 177  
 QY 208 GYDTKSDIWSLACILYEMCCMHAFAGSNFLSIVLKIIVEGDTPLSPRYPKELNAMESM 267  
 DB 178 PYNKSDIWSLGCVLIELCTLKHFPQANSWKNILKVCQCISPLFSHYSELOFLVKQM 237  
 QY 268 LNKNPSPRPSAIEILKIPYDEQLQ---NLMCRYSEMTELED-KNL-----DC 310  
 DB 238 FKRNPSHRPSATLLSRGIVARLVQKLPPEIIMEYGEVLEIKNSKNTPRKQDEBEQ 297  
 QY 311 OKERAAHI-INAMOKRIHLQTLRALSEVQKMTPEERVRLKLAQADEKARKKIVVE--- 366  
 DB 298 DRKGSHTDLESINENLVESALRRVNEEK---GNKSVHLRRKASSPNLHRKQWKNQVNTAL 355  
 QY 367 -----TALNASILTSLSAETDRGSGVIKYSKNTTRKQWLKETPDTLLNLKNVADLSLAFQTV 415  
 DB 356 TDVNLHEKTHLKG---MEEKES-----OPEGRILSCSPQDEERWQGRREESDE 432  
 QY 416 IYRPGSEGFLKGPLSEETASDSVDCGHSDVILDPE-RLEPGLDEEDTD-----FEEEDN 470

Search completed: October 1, 2004, 11:20:45.  
Job time : 125 secs

```
OY 433 P 433
Db 471 P 471

RESULT 15
QBBSB6
ID QBBSB6 PRELIMINARY; PRT; 302 AA.
AC QBBSB6;
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE NIMA.
GN NEKI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=22354693; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573(2002).
DR EMBL; AK034754; BAC28822.1; -.
DR MGD; MGI:97303; Neki.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; P:protein-tyrosine kinase activity; IEA.
DR GO; GO:0004688; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_kin_AS.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE_ST; 1.
SQ SEQUENCE 302 AA; 34328 MW; 6D5B14D7809CB341 CRC64;

Query Match 23.1%; Score 577; DB 11; Length 302;
Best Local Similarity 38.5%; Pred.No. 4.7e-31;
Matches 110; Conservative 71; Mismatches 95; Indels 10; Gaps 5;

OY 28 RYVLQKLGSGGTGYLYSDKKRGEELVYKETSVCGLNPNETVOANLEAQLLSKLD 87
Db 3 KYVRLQIGEGSGKAVLV--KSTEDGRHY-VIKEINISRMDSKQESRRVAVLANWK 59

OY 88 HPAIVKPHASFEVDNFCITTEYCEGRDLDDKIOEYKQAGKIPPENQIIWFQILLGVD 147
Db 60 HPNIVQKESFEENGSLYIVMDYCEGDLFXRN--AQKALFQEDQILDWFVQICLALK 117

OY 148 YMHERRILHDLKSKVFL-KNNLLKIGDGVSRLLMGSCDLATTLTGTPHYMSPEALKH 206
Db 118 HVHDKRILHDKSQNIFLTQDGTQVQDGFARVLNSTVELARTCTIGTPYLSPEICEN 177

OY 207 QGYDTKDSIWSLACILYVCMCNHAFAGNSFLSVLKIIVEGDTPTSLPERYPKELNAIMES 266
Db 178 KPYNNKSDIHALGCVLYELCTLKHAEAGNMKNLVLIKISGSPFPVSPHYSDLRSLLSQ 237

OY 267 MLNKNPSLRPSAIEILKIPYLDQLQNL-----CRYSEMTLEDKNL 308
Db 238 LFKRNPDRPNSVNSILEKFIKRIEKLSPQVCVTLPLTLCKVL 283
```

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2004, 11:43:43 ; Search time 5789 Seconds  
(without alignments)  
10848.861 Million cell updates/sec

Title: US-10-803-278-3

Perfect score: 1449  
Sequence: 1 atgtcgaaatcccaagagc.....actggaaaaatgaaaaataa 1449

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:\*

1: gb\_ba:\*

2: gb\_htg:\*

3: gb\_in:\*

4: gb\_om:\*

5: gb\_ov:\*

6: gb\_pat:\*

7: gb\_ph:\*

8: gb\_pl:\*

9: gb\_pr:\*

10: gb\_ro:\*

11: gb\_scs:\*

12: gb\_sy:\*

13: gb\_un:\*

14: gb\_vl:\*

15: em\_ba:\*

16: em\_fun:\*

17: em\_hum:\*

18: em\_in:\*

19: em\_mu:\*

20: em\_om:\*

21: em\_or:\*

22: em\_ov:\*

23: em\_pat:\*

24: em\_ph:\*

25: em\_pl:\*

26: em\_ro:\*

27: em\_scs:\*

28: em\_un:\*

29: em\_vl:\*

30: em\_htg\_hum:\*

31: em\_htg\_inv:\*

32: em\_htg\_other:\*

33: em\_htg\_mus:\*

34: em\_htg\_pln:\*

35: em\_htg\_rod:\*

36: em\_htg\_mam:\*

37: em\_htg\_vrt:\*

38: em\_sy:\*

39: em\_htgo\_hum:\*

40: em\_htgo\_mus:\*

41: em\_htgo\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1449	100.0	1854	9	BC028587 Homo sapi
2	1399.8	96.6	2080	9	AB071997 Homo sapi
3	1399	96.5	1938	6	AX166555 Sequence
4	1399	96.5	2552	6	AX746174 Sequence
5	1399	96.5	2768	9	HS0604785
6	1399	96.5	2939	9	AB071996 Homo sapi
7	1335	92.1	2936	9	AB064997
8	918	63.4	2049	9	AK027148 Homo sapi
9	912.2	63.0	2902	9	AB066559 Macaca fa
10	651.8	45.0	2189	9	AK123021 Homo sapi
11	172.2	11.9	154330	2	AC025148
12	172.2	11.9	133425	9	AC055733
13	167.2	11.5	60398	9	AC116424 Homo sapi
14	153.4	10.6	111568	9	AC083908 Homo sapi
15	153.4	10.6	175923	2	AC064839 Homo sapi
16	152.2	10.5	1547	3	AK114791 Ciona int
17	148	10.2	2603	9	BC063044 Homo sapi
18	148	10.2	3698	9	HUMSTK2A
19	147.8	10.2	1510	6	AX657954 Sequence
20	147.8	10.2	1578	6	AX657957
21	147.8	10.2	1581	6	AX657962 Sequence
22	147.8	10.2	1610	9	AB072828 Homo sapi
23	147.8	10.2	1638	6	AX803387 Sequence
24	147.8	10.2	2059	6	BD251812 Phosphory
25	147.8	10.2	2110	6	AX504240 Sequence
26	147.8	10.2	2134	9	BC019916 Homo sapi
27	146.4	10.1	2463	10	MMSTRK2S
28	146.4	10.1	2607	10	MMSTRK2L
29	146.4	10.1	2715	6	AX657960 Sequence
30	146.4	10.1	2715	6	AX659642
31	146.4	10.1	2715	10	AF099057
32	146.4	10.1	2778	10	BC057939
33	146.2	10.1	1868	9	HSNEK3R
34	146.2	10.1	2364	6	AX405633 Sequence
35	146.2	10.1	4263	6	AR223879 Sequence
36	146.2	10.1	4263	10	S45828
37	143.2	9.9	3645	6	AX224471
38	143.2	9.9	3729	6	AX658011 Sequence
39	143.2	9.9	5426	6	AX224518
40	143.2	9.9	5448	6	AR338755
41	143.2	9.9	5497	9	AB067488
42	143.2	9.9	5508	6	AX658014
43	143.2	9.9	5511	9	HS0800531
44	143.2	9.9	5532	6	AR338754
45	140.4	9.7	1997	10	AF093416 Mus muscu

# ALIGNMENTS

RESULT 1  
BC028587  
LOCUS BC028587 1854 bp mRNA linear PRI 07-OCT-2003  
DEFINITION Homo sapiens NIMA (never in mitosis gene a) - related kinase 11,  
mRNA (CDNA clone MGC:27011 IMAGE:4829802), complete cds.  
ACCESSION BC028587  
VERSION BC028587.2 GI:34191901  
KEYWORDS MGC  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1854)  
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,  
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,

Pred. No. is the number of results predicted by chance to have a

Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schetz, T.B., Brownstein, M.J., Usdin, T.B., Tohivuyki, S., Carninci, P., Frange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S.E., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.N., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.C., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalhus, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2238257

12477932

2 (bases 1 to 1854)

Strausberg, R.

Direct Submission

Submitted (23-APR-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

On Aug 25, 2003 this sequence version replaced gi:20306789.

Contact: MGC help desk

Email: [cgaabs@mail.nih.gov](mailto:cgaabs@mail.nih.gov)

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.

cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki

Toshiyuki and Piero Carninci (RIKEN)

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Institute for Systems Biology

<http://www.systemsbio.org>

contact: [amadnan@systemsbiology.org](mailto:amadnan@systemsbiology.org)

Anup Madan, Jessica Fahy, Erin Helton, Mark Kettman, Anuradha

Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 34 Row: 9 Column: 23

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 22094142.

#### FEATURES

##### source

1. .1854  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="MGC:27011 IMAGE:4829802"  
/tissue\_type="testis"  
/clone\_lib="NIH\_MGC\_97"  
/lab\_host="DH10B"  
/note="Vector: pBluescript"

##### gene

1. .1854  
/gene="NEK11"  
/note="synonym: FLJ23495"  
/db\_xref="LocusID:79858"  
371..1819  
/codon\_start=1  
/product="NEK11 protein"  
/protein\_id="AAH28587.1"  
/db\_xref="GI:20306790"

##### CDS

/translation="MLKFEAAKCVSGTAISTYPTLIARRYLOQLGSGFGTVY  
LVSDKAKGEBELKVLKEISVGLNPNETVOANLEQLSLKLDHPAIKVFHASFVEQD  
NFCITTEYCEGRLDDKIDIEYKQAGKIPFENQIIEWFQILLGVDMERILLRDLK  
SKNVLKNNLKIIGDFVGRLLMGCDLATTUTGPHYNSPEALHGQIDTKSDIWSL  
ACTLYLKMCMNHAFAAGSLVSLVILKIVEGDTSPERPYPKELNAMESMLKNPSLRP  
SAIEILKIPYLDELQNLNCRYSEMTLEDKLDCKQEAHHIINAMQRIHLQTLRALS

EVQVQTPRRMBRLKLOADEKARKLKIIVEEKYEENSKMOELSRNFQQLSVDLVH  
EKTHLKGMEKEBEPQERLSCSPQEDERWQGRREESDEPLENLPSPQIPPSMDLH  
ELESIVEDATSLGVHGDNCNLISLDEYWNKEK"  
455..1231  
misc\_feature  
/note="S TKC; Region: Serine/Threonine protein kinases,  
catalytic domain"  
/db\_xref="CDD:smart00220"

#### ORIGIN

Query Match 100.0%; Score 1449; DB 9; Length 1854;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1449; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATCTGAAATTCACAGAGCAGCTAAGTGTGTGAGTGGATCAACAGCCATTTCCTTAT 60  
Db 371 ATCTGAAATTCACAGAGCAGCTAAGTGTGTGAGTGGATCAACAGCCATTTCCTTAT 430  
Qy 61 CCAAGACCTTGATTGCAAGAAGATACGTCTTCAACAAAACTTGGCAGTGGAGTTT 120  
Db 431 CCAAGACCTTGATTGCAAGAAGATACGTCTTCAACAAAACTTGGCAGTGGAGTTT 490  
Qy 121 GGAAGTGTCTATCTGTTTTCAGACACAGAACCCACAGGAGGAGGAATTAAGGTACTT 180  
Db 491 GGAAGTGTCTATCTGTTTTCAGACACAGAACCCACAGGAGGAGGAATTAAGGTACTT 550  
Qy 181 AAGGAATATCTCTTGGAGAACTTAAATCCAAATGAAACTGTACAGCCAAATTTGGAAGCC 240  
Db 551 AAGGAATATCTCTTGGAGAACTTAAATCCAAATGAAACTGTACAGCCAAATTTGGAAGCC 610  
Qy 241 CAATCTCTTCAAGCTGGACACACAGCCATTGTCAAGTTCATGCAAGTTCCTTGTGGAG 300  
Db 611 CAATCTCTTCAAGCTGGACACACAGCCATTGTCAAGTTCATGCAAGTTCCTTGTGGAG 670  
Qy 301 CAAGATAATTTCTGCATTATCACGAGTACTGTGAGGCGCGAGATCTGGACGATTAAT 360  
Db 671 CAAGATAATTTCTGCATTATCACGAGTACTGTGAGGCGCGAGATCTGGACGATTAAT 730  
Qy 361 CAGGAATATAAACAAGCTGAAAAATCTTTCCGAAAAATCAATAATAGATGGTTTATC 420  
Db 731 CAGGAATATAAACAAGCTGAAAAATCTTTCCGAAAAATCAATAATAGATGGTTTATC 790  
Qy 421 CAGTCTGCTGGAGTGTACTACATGATGAGGAGGATACCTTCATCGAGACTTTAAAG 480  
Db 791 CAGTCTGCTGGAGTGTACTACATGATGAGGAGGATACCTTCATCGAGACTTTAAAG 850  
Qy 481 TCAAGAAATGATTTCTGAAAAATAATCTCTTAAATTTGGAGATTTGGAGTTTTCGA 540  
Db 851 TCAAGAAATGATTTCTGAAAAATAATCTCTTAAATTTGGAGATTTGGAGTTTTCGA 910  
Qy 541 CTTCATAGGATCTGTGACCTGGCCCAACACTTAACTGGACTCCCATATATAGT 600  
Db 911 CTTCATAGGATCTGTGACCTGGCCCAACACTTAACTGGACTCCCATATATAGT 970  
Qy 601 CTTGAGGCTCTGAAACACCAAGGCTATGACACAAAGTCGACATCTGTCATGGCATGC 660  
Db 971 CTTGAGGCTCTGAAACACCAAGGCTATGACACAAAGTCGACATCTGTCATGGCATGC 1030  
Qy 661 ATTTTGTATGAGATGTCTGCATGATCATGATTCGCTGGCTCCCAATTTCTTATCCATT 720  
Db 1031 ATTTTGTATGAGATGTCTGCATGATCATGATTCGCTGGCTCCCAATTTCTTATCCATT 1090  
Qy 721 GTTTTAAAAATTTGTTGAAGGTGACACACCTTCTCTCTCCCTGAGAGATATCCAAAGAACTA 780  
Db 1091 GTTTTAAAAATTTGTTGAAGGTGACACACCTTCTCTCTCCCTGAGAGATATCCAAAGAACTA 1150  
Qy 781 AATGCCATCATGGAAGCATGTTGAAACAAGATCTTTCATTAAGACCATCTGCTATCGAA 840  
Db 1151 AATGCCATCATGGAAGCATGTTGAAACAAGATCTTTCATTAAGACCATCTGCTATCGAA 1210  
Qy 841 ATTTTAAAAATTCCTTACCTTGTGATGAGCAGTACAGAACCTAATGTGTAGATATTAGAA 900  
Db 1211 ATTTTAAAAATTCCTTACCTTGTGATGAGCAGTACAGAACCTAATGTGTAGATATTAGAA 1270



```
QY 901 ATGACTCTGGAAGACAAAATTTGGATTGTCAGAGGAGGCTGCTCATATATTAATGCC 960
Db 1271 ATGACTCTGGAAGACAAAATTTGGATTGTCAGAGGAGGCTGCTCATATATTAATGCC 1330
QY 961 ATCAAAAAAGGATCCACCTGCAGACTCTGAGGGCACTGTCAGAGTACAGAAAAATGAGC 1020
Db 1331 ATCAAAAAAGGATCCACCTGCAGACTCTGAGGGCACTGTCAGAGTACAGAAAAATGAGC 1390
QY 1021 CCAAGAGAAAGATGCGGCTGAGGAAGCTCCAGCGGCTGATGAGAAACCCAGGAAGCTG 1080
Db 1391 CCAAGAGAAAGATGCGGCTGAGGAAGCTCCAGCGGCTGATGAGAAACCCAGGAAGCTG 1450
QY 1081 AAAAAAGATTGTGAAGAAAAATATGAAGAAATAGCAAAACGAATGCAAGAAATTGAGATCT 1140
Db 1451 AAAAAAGATTGTGAAGAAAAATATGAAGAAATAGCAAAACGAATGCAAGAAATTGAGATCT 1510
QY 1141 CGGAATTTTCAGCAGCTGATGTTGATGTAATCTCCATGAAACCAATTTTAAAGGAATG 1200
Db 1511 CGGAATTTTCAGCAGCTGATGTTGATGTAATCTCCATGAAACCAATTTTAAAGGAATG 1570
QY 1201 GAAGAAAAAGGAGGACCACTGAGGGAAGACTTTCTGTTTCAACCCAGGACGAGGATGAA 1260
Db 1571 GAAGAAAAAGGAGGACCACTGAGGGAAGACTTTCTGTTTCAACCCAGGACGAGGATGAA 1630
QY 1261 GAGAGTGCAAGGCAAGGGAAGGAATCTGATGAACCAACTTTAGAGAACCTGCTGAG 1320
Db 1631 GAGAGTGCAAGGCAAGGGAAGGAATCTGATGAACCAACTTTAGAGAACCTGCTGAG 1690
QY 1321 TCTCAGCTTATCTCTTCATGAGCTCCAGGACCTGAGAACTGATGTAAGTGAAGTGAAGTGA 1380
Db 1691 TCTCAGCTTATCTCTTCATGAGCTCCAGGACCTGAGAACTGATGTAAGTGAAGTGAAGTGA 1750
QY 1381 TCTGACCTTGGATACCATGAGGACTGTAATCTTAATTTCACTAGAGAACTACTGGAATAAT 1440
Db 1751 TCTGACCTTGGATACCATGAGGACTGTAATCTTAATTTCACTAGAGAACTACTGGAATAAT 1810
QY 1441 GAAAAATAA 1449
Db 1811 GAAAAATAA 1819

RESULT 2
LOCUS AB071997 2080 bp mRNA linear PRI 17-OCT-2002
DEFINITION Homo sapiens nek1s mRNA for NIMA-related kinase 11S, complete cds.
ACCESSION AB071997
VERSION AB071997.1 GI:21955953
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Noguchi,K., Fukazawa,H., Murakami,Y. and Uehara,Y.
TITLE Nek1, a new member of the NIMA family of kinases, involved in DNA
JOURNAL replication and genotoxic stress responses
MEDLINE J. Biol. Chem. 277 (42), 39555-39565 (2002)
PUBMED 22267232
REFERENCE 2 (bases 1 to 2080)
AUTHORS Noguchi,K., Uehara,Y., Fukazawa,H. and Murakami,Y.
TITLE Direct Submission
JOURNAL Submitted (25-SEP-2001) Kohji Noguchi, National Institute of
Infectious Diseases, Department of Bioactive Molecules; 1-23-1
Toyama, Shinjuku-ku, Tokyo 132-0023, Japan
(E-mail:knoguchi@nih.go.jp, Tel:81-3-5285-1111 (ex.2327),
Fax:81-3-5285-1272)
FEATURES
Location/Qualifiers
1..2080
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="3q21-q22"
```

```
/cell_lines="HeLaS3"
1..2080
/genes="nek1s"
294..1706
/genes="nek1s"
/codon_start=1
/product="NIMA-related kinase 11S"
/protein_id="BAC06351.1"
/db_xref="GI:21955954"
/translation="MLKSFQAAKCVSGSTAISTYPTKLTIAIRVYLQKLGSGSGFTVY
LVSKAKRGEEELXKEISVGLNPNETVQANLEAQLLSKLDHPAIVKPHASVEQD
NFCITTEYGRDLDKIOEVKOAGKIPFNQIIEWFIQLLGVYVHERRILHRDLK
SKNVLKNLLKIGDFVSRLLMGSCDLATTLTGTHVMSPEALKHOGYDTKSPDMSI
ACILYECMWHAFAGSNFLSIVLKIIVEGTPSPERYPKELNAMESMLNKPSPSLP
SAISILKIPYDEQLQNLKYSKTEIDKNLDCQKRAHLINWQKRIHQITLRAUS
EVQNTPREMRLEKQADEKARKLKIIVEKYENSKRMQBLRSNFOLSVDDVLH
EKTLKGMSEKEQEGRLSCQDEDEERWQREESDEFTLENLPESQIPISMDLH
ELESIVEDATSDLYHATHS"
ORIGIN
Query Match 96.6%; Score 1399.8; DB 9; Length 2080;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1407; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 1 ATGCTGAAATCCAGAGGAGCAGCTAAGTGTGTGAGTGGATCAACAGACCAATTCACATTAT 60
Db 294 ATGCTGAAATCCAGAGGAGCAGCTAAGTGTGTGAGTGGATCAACAGACCAATTCACATTAT 353
QY 61 CCAAGAGCCTTGATGCAAGAGATACGTGCTTCAACAAAAAATTGGCAGTGGAAAGTTT 120
Db 354 CCAAGAGCCTTGATGCAAGAGATACGTGCTTCAACAAAAAATTGGCAGTGGAAAGTTT 413
QY 121 GGAAGTGTCTATCTGTTTTCAGACAAAGAAAGCCAAAGAGGAGAGGAATTAAGGTACTT 180
Db 414 GGAAGTGTCTATCTGTTTTCAGACAAAGAAAGCCAAAGAGGAGAGGAATTAAGGTACTT 473
QY 181 AAGAAATATCTGTTGGAGAACTAAATCCAAATGAAACTGTACAGGCCAATTTGGAAGCC 240
Db 474 AAGAAATATCTGTTGGAGAACTAAATCCAAATGAAACTGTACAGGCCAATTTGGAAGCC 533
QY 241 CAAGTCTCTCCAGCTGGACCCAGCCAGCCATGTCAGGTTCCATGCAAGTTTGTGGAG 300
Db 534 CAAGTCTCTCCAGCTGGACCCAGCCAGCCATGTCAGGTTCCATGCAAGTTTGTGGAG 593
QY 301 CAAGATAATTTCTGATTATCACGGAGTACTGTGAGGCGCCGAGATCTGGACGATAAAAT 360
Db 594 CAAGATAATTTCTGATTATCACGGAGTACTGTGAGGCGCCGAGATCTGGACGATAAAAT 653
QY 361 CAGGAATATAACAGAGCTGGAAAAATCTTTCCAGAAAAATCAATATAGATGGTTTATC 420
Db 654 CAGGAATATAACAGAGCTGGAAAAATCTTTCCAGAAAAATCAATATAGATGGTTTATC 713
QY 421 CAGCTGCTGCTGGAGTTGACTACATCATGAGAGGAGGATCTTTCATCGAGACTTAAAG 480
Db 714 CAGCTGCTGCTGGAGTTGACTACATCATGAGAGGAGGATCTTTCATCGAGACTTAAAG 773
QY 481 TCAAGAAATGTATTTCTGAAAAAATAATCTCTTAAATTTGAGATTTTGAGTTTCTCGA 540
Db 774 TCAAGAAATGTATTTCTGAAAAAATAATCTCTTAAATTTGAGATTTTGAGTTTCTCGA 833
QY 541 CTTCCTAATGGATCCTGTGACCTGGCCACAACTTTTAAGTGGAACTCCCATTTATATGAGT 600
Db 834 CTTCCTAATGGATCCTGTGACCTGGCCACAACTTTTAAGTGGAACTCCCATTTATATGAGT 893
QY 601 CCTGAGGCTCTGAACACCAAGGCTATGACACAAAGTCGAGACATCTGGTCACTGGCATGC 660
Db 894 CCTGAGGCTCTGAACACCAAGGCTATGACACAAAGTCGAGACATCTGGTCACTGGCATGC 953
QY 661 ATTTTGTATGAGTGTCTGATGAATCATGCAATTCGCTGGCTCCATTTCTTATCCATT 720
Db 954 ATTTTGTATGAGTGTCTGATGAATCATGCAATTCGCTGGCTCCATTTCTTATCCATT 1013
QY 721 GTTTTAAAAAATTTGTTGAAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAAAGACTA 780
```

Db 1014 GTTTTAAATTTGTTGAAGTGACACACACTTCTCTCCCTGAGAGATATCCAAAGAACTA 1073  
QY 781 AATGCCATCATGAAAGCATGTTGAACAGAAATCCTTCAATTAAGACCATCTGCTATCGAA 840  
Db 1074 AATGCCATCATGAAAGCATGTTGAACAGAAATCCTTCAATTAAGACCATCTGCTATCGAA 1133  
QY 841 ATTTTAAATTCCTTACCTTGTATGAGCAGCTACAGAACTTAATGTGTAGATATTTCAGAA 900  
Db 1134 ATTTTAAATTCCTTACCTTGTATGAGCAGCTACAGAACTTAATGTGTAGATATTTCAGAA 1193  
QY 901 ATGACTCTGGAAGACAAAATTTGGATTGTGAGAGGAGCTGCTCATATAATTAATGCC 960  
Db 1194 ATGACTCTGGAAGACAAAATTTGGATTGTGAGAGGAGCTGCTCATATAATTAATGCC 1253  
QY 961 ATGCAAAAAGGATCCACCTGAGACTCTGAGGGCACTGTGAGAGTACAGAAATGAGC 1020  
Db 1254 ATGCAAAAAGGATCCACCTGAGACTCTGAGGGCACTGTGAGAGTACAGAAATGAGC 1313  
QY 1021 CCAAGAGAAAGGATCGGGCTGAGGAAGCTCCAGGGCTGTGAGAAAGCCAGGAAGCTG 1080  
Db 1314 CCAAGAGAAAGGATCGGGCTGAGGAAGCTCCAGGGCTGTGAGAAAGCCAGGAAGCTG 1373  
QY 1081 AAAAGGATCTGCAAGAAATAATGAGAAATAGCAACGATGCAAGAAATGAGATCT 1140  
Db 1374 AAAAGGATCTGCAAGAAATAATGAGAAATAGCAACGATGCAAGAAATGAGATCT 1433  
QY 1141 CGGAATCTTCAGCAGCTGAGTGTGTGATGTACTCCATGAAAAACACATTTAAAGGAATG 1200  
Db 1434 CGGAATCTTCAGCAGCTGAGTGTGTGATGTACTCCATGAAAAACACATTTAAAGGAATG 1493  
QY 1201 GAAGAAAGAGAGAGCAACTGAGGAAGACTTCTTGTGACCCAGCAGAGAGTGA 1260  
Db 1494 GAAGAAAGAGAGAGCAACTGAGGAAGACTTCTTGTGACCCAGCAGAGAGTGA 1553  
QY 1261 GAGAGTGGCAAGCAGGAGAGGAATCTGTGAACCAACTTTAGAGAACTGCTGAG 1320  
Db 1554 GAGAGTGGCAAGCAGGAGAGGAATCTGTGAACCAACTTTAGAGAACTGCTGAG 1613  
QY 1321 TCTCAGCTATTCTCTCCATGAGACCTCCAGAACTTGAATCAATTTGAGAGATGCCACA 1380  
Db 1614 TCTCAGCTATTCTCTCCATGAGACCTCCAGAACTTGAATCAATTTGAGAGATGCCACA 1673  
QY 1381 TCTGACCTTGATACCATGAGAGCTGTAATCTAATTCA 1419  
Db 1674 TCTGACCTTGATACCATGAGAGCTGTAATCTAATTCA 1712

## RESULT 3

AX166555

LOCUS

AX166555

DEFINITION

Sequence 46 from Patent WO0138503.

ACCESSION

AX166555

VERSION

AX166555.1

KEYWORDS

GI:14546900

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1

Plowman, G.D., Whyte, D., Manning, G.S., Sudarsanam, S., Martinez, R.,

Flanagan, P., and Clary, D.S.

Novel human protein kinases and protein kinase-like enzymes

Patent: WO 0138503-A 46 31-MAY-2001;

Sugen, Inc. (US)

FEATURES

Location/Qualifiers

1..1938

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

## ORIGIN

Query Match

96.5%; Score 1399; DB 6; Length 1938;

Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTGAAATTCGAAGAGGAGCTAAGTGTGTGAGTGAATCAACAGCATTTCCTCAT 60  
Db 1 ATGCTGAAATTCGAAGAGGAGCTAAGTGTGTGAGTGAATCAACAGCATTTCCTCAT 60  
QY 61 CCAAGACCTTGATGTCGAAGAGATACGTCCTTCAACAAAACTTGGCAGTGAAGTTT 120  
Db 61 CCAAGACCTTGATGTCGAAGAGATACGTCCTTCAACAAAACTTGGCAGTGAAGTTT 120  
QY 121 GGAATCTCTATCTGGTTTCAGACAGAGAGCCAAACGAGAGAGAAATTAAGAGTACTT 180  
Db 121 GGAATCTCTATCTGGTTTCAGACAGAGAGCCAAACGAGAGAGAAATTAAGAGTACTT 180  
QY 181 AAGGAAATATCTGTGTGAGAACTAAATCCAAATCTGACAGGCCAAATTTGGAAGCC 240  
Db 181 AAGGAAATATCTGTGTGAGAACTAAATCCAAATCTGACAGGCCAAATTTGGAAGCC 240  
QY 241 CAATCTCTTCCAAGCTGGACCCAGCCATTGTCAAGTTCATGCAAGTTTGTGGAG 300  
Db 241 CAATCTCTTCCAAGCTGGACCCAGCCATTGTCAAGTTCATGCAAGTTTGTGGAG 300  
QY 301 CAAGATAATTTCTGCATTATCAGGAGTACTGTGAGGCGCGAGATCTGGACGATAAAT 360  
Db 301 CAAGATAATTTCTGCATTATCAGGAGTACTGTGAGGCGCGAGATCTGGACGATAAAT 360  
QY 361 CAGAAATATAAACAAGCTGGAAAAATCTTTCAGAAAAATCAAAATATAGATGGTTATC 420  
Db 361 CAGAAATATAAACAAGCTGGAAAAATCTTTCAGAAAAATCAAAATATAGATGGTTATC 420  
QY 421 CAGCTGCTGTGGAGTTGACTACATGATGAGAGAGGATCTTCATCGAGACTTAAAG 480  
Db 421 CAGCTGCTGTGGAGTTGACTACATGATGAGAGAGGATCTTCATCGAGACTTAAAG 480  
QY 481 TCAAGAAATGTATTTCTGAAAAATTAATCTCCTTAAAAATTTGGAGTTTCTCGA 540  
Db 481 TCAAGAAATGTATTTCTGAAAAATTAATCTCCTTAAAAATTTGGAGTTTCTCGA 540  
QY 541 CTTCTAATGGATCCTGTGACCTGGCCACAACTTAACTGGAATCCCAATATATAGT 600  
Db 541 CTTCTAATGGATCCTGTGACCTGGCCACAACTTAACTGGAATCCCAATATATAGT 600  
QY 601 CTTGAGGCTCTGAAACACCAAGGCTATGACACAAAGTCCGACATCTGGTCACTGGCATGC 660  
Db 601 CTTGAGGCTCTGAAACACCAAGGCTATGACACAAAGTCCGACATCTGGTCACTGGCATGC 660  
QY 661 ATTTTGTATGAGATGCTGCTCATGAAATCATGCAATTCGCTGGCTCCATTTCTTATCCATT 720  
Db 661 ATTTTGTATGAGATGCTGCTCATGAAATCATGCAATTCGCTGGCTCCATTTCTTATCCATT 720  
QY 721 GTTTTAAAAATTTGTTGAAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTA 780  
Db 721 GTTTTAAAAATTTGTTGAAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTA 780  
QY 781 AATGCCATCATGAAAGCATGTTGAACAGAAATCCTTCAATTAAGACCATCTGCTATCGAA 840  
Db 781 AATGCCATCATGAAAGCATGTTGAACAGAAATCCTTCAATTAAGACCATCTGCTATCGAA 840  
QY 841 ATTTTAAAAATTCCTTACCTTGTATGAGAGGAGCTGCTCATATAATTAATGCC 900  
Db 841 ATTTTAAAAATTCCTTACCTTGTATGAGAGGAGCTGCTCATATAATTAATGCC 900  
QY 901 ATGACTCTGGAAGACAAAATTTGGATTGTGAGAGGAGCTGCTCATATAATTAATGCC 960  
Db 901 ATGACTCTGGAAGACAAAATTTGGATTGTGAGAGGAGCTGCTCATATAATTAATGCC 960  
QY 961 ATGCAAAAAGGATCCACCTGAGACTCTGAGGGCACTGTGAGAGTACAGAAATGAGC 1020  
Db 961 ATGCAAAAAGGATCCACCTGAGACTCTGAGGGCACTGTGAGAGTACAGAAATGAGC 1020  
QY 1021 CCAAGAGAAAGGATCGGGCTGAGGAAGCTCCAGGGCTGTGAGAAAGCCAGGAAGCTG 1080

```
Db 1021 CCAAGAGAAAGGATCGGCTGAGGAAGCTCCAGGCGGCTGATGAGAAAGCCAGGAAGCTG 1080
Qy 1081 AAAAGATTGTGGAAGAAAATATCAAGAAAATAGCAAGAAATGCAAGAAATGAGATCT 1140
Db 1081 AAAAGATTGTGGAAGAAAATATGAGAAAATAGCAAGAAATGCAAGAAATGAGATCT 1140
Qy 1141 CGAACTTTACAGCAGCTGAGTGTGATGTACTCCATGAAAGAAACACATTTAAAGGAATG 1200
Db 1141 CGAACTTTACAGCAGCTGAGTGTGATGTACTCCATGAAAGAAACACATTTAAAGGAATG 1200
Qy 1201 GAAGAAAGGAGGACCACTGAGGAGAGACTTCTGTTTCAACCCAGGAGCGAGGATGAA 1260
Db 1201 GAAGAAAGGAGGAGCAACTGAGGAGAGACTTCTGTTTCAACCCAGGAGCGAGGATGAA 1260
Qy 1261 GAGAGGTGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
Db 1261 GAGAGGTGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
Qy 1321 TCTGAGCTATTCTTCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
Db 1321 TCTGAGCTATTCTTCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
Qy 1381 TCTGAGCTATTCTTCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1399
Db 1381 TCTGAGCTATTCTTCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1399

RESULT 4
LOCUS AX746174 2552 bp DNA linear PAT 13-JUN-2003
DEFINITION Sequence 25 from Patent WO0208399.
ACCESSION AX746174
VERSION AX746174.1 GI:31746160
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yue,H., Khan,F.A., Gururajan,R., Hafalia,A.J., Chawla,N.K.,
Arvizu,C.S., Ramkumar,J., Gandhi,A.R., Policky,J.L., Baughn,M.R.,
Tribouley,C.M., Bandman,O., Nguyen,D.B., Lu,Y., Burford,N., Lal,P.,
Ding,L., Yao,M.G., Elliott,V.S., Recipon,S.A., Kearney,L., Lu,D.A.,
Greenwald,S.R., Tang,Y.T., Xu,Y., Walsh,R.T., Gietzen,K.J.,
Yang,J., Jackson,J.L. and Thornton,M.
HUMAN KINASES
TITLE Patent: WO 0208399-A 25 31-JAN-2002;
JOURNAL Incyte Genomics, Inc. (US) ; Thornton, Michael (US)
FEATURES
Location/Qualifiers
1..2552
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7477585CB1"
ORIGIN
Query Match 96.5%; Score 1399; DB 6; Length 2552;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGCTGAATTCAGAGCAGCAGTAAAGTGTGTGAGTGGATCAACAGCAATTTCCACTTAT 60
Db 278 ATGCTGAATTCAGAGCAGCAGTAAAGTGTGTGAGTGGATCAACAGCAATTTCCACTTAT 337
Qy 61 CCAAGACCTTGATTGCAAGAGATACGCTTCAACAAAACCTTGGCAGTGGAGTTT 120
Db 338 CCAAGACCTTGATTGCAAGAGATACGCTTCAACAAAACCTTGGCAGTGGAGTTT 397
Qy 121 GGAAGTGTCTATCTGTTTTCAGACAGAAAGCCAAACGAGGAGAGGAAATTAAGGTACT 180
Db 398 GGAAGTGTCTATCTGTTTTCAGACAGAAAGCCAAACGAGGAGAGGAAATTAAGGTACT 457
Qy 181 AGGAAATATCTGTTGGAGAACTAAATCCAAATGAACTGTACAGGCCAATTTGGAAGCC 240
```

```
Db 458 AAGGAATATCTGTGGAGAACTAAATCCAAATGAAATCTACAGGCCAATTTGGAGCC 517
Qy 241 CAATCTCTCTCCAAAGCTGGACCAACCCAGCCATTTGCAAGTTCCATGCAAGCTTTTGTGGAG 300
Db 518 CAATCTCTCTCCAAAGCTGGACCAACCCAGCCATTTGCAAGTTCCATGCAAGCTTTTGTGGAG 577
Qy 301 CAAGATAAATTTCTGCATTTATCAGGAGTACTGTGAGGCGCGAGATCTGGACGATATAAATT 360
Db 578 CAAGATAAATTTCTGCATTTATCAGGAGTACTGTGAGGCGCGAGATCTGGACGATATAAATT 637
Qy 361 CAGGAATATAAACAAGCTGGAATAATCTTCCAGAAAATCAATATAATAGATGTTTATC 420
Db 638 CAGGAATATAAACAAGCTGGAATAATCTTCCAGAAAATCAATATAATAGATGTTTATC 697
Qy 421 CAGCTGCTGTGGAGTTGACTACATGATGAGAGGAGGATCTTTCATCGAGACTTTAAAG 480
Db 698 CAGCTGCTGTGGAGTTGACTACATGATGAGAGGAGGATCTTTCATCGAGACTTTAAAG 757
Qy 481 TCAAGAAATGATTTCTGAAAATAATCTCTTAAATAATGGAGATTTTGGAGTTTCTCGA 540
Db 758 TCAAGAAATGATTTCTGAAAATAATCTCTTAAATAATGGAGATTTTGGAGTTTCTCGA 817
Qy 541 CTTCTAATGGATCTCTGACCTGGCCACAACTTTAACTGAACTCCCAATTATATGAGT 600
Db 818 CTTCTAATGGATCTCTGACCTGGCCACAACTTTAACTGAACTCCCAATTATATGAGT 877
Qy 601 CTTGAGGCTCTGAAAACAACCAAGGCTATGACAAAAGTCGACATCTGGTCACTGGCATGC 660
Db 878 CTTGAGGCTCTGAAAACAACCAAGGCTATGACAAAAGTCGACATCTGGTCACTGGCATGC 937
Qy 661 ATTTTGTATGAGATGTGCTGCATGATCATGCAATTCGCTGGCTCCCAATTTCTTATCCATT 720
Db 938 ATTTTGTATGAGATGTGCTGCATGATCATGCAATTCGCTGGCTCCCAATTTCTTATCCATT 997
Qy 721 GTTTTAAAAATTTGTGAAGGTGACACACCTTTCTCTCCCTGAGAGATATCAAAAGAACTA 780
Db 998 GTTTTAAAAATTTGTGAAGGTGACACACCTTTCTCTCCCTGAGAGATATCAAAAGAACTA 1057
Qy 781 AATGCCATCATGGAAGCATGTTGAAACAAGATCCTTCAATAGACCATCTGCTATCGAA 840
Db 1058 AATGCCATCATGGAAGCATGTTGAAACAAGATCCTTCAATAGACCATCTGCTATCGAA 1117
Qy 841 ATTTTAAAAATTCCTTTACCTTTGATGACAGCTACAGAACCTTAATGTGTAGATATTCAGAA 900
Db 1118 ATTTTAAAAATTCCTTTACCTTTGATGACAGCTACAGAACCTTAATGTGTAGATATTCAGAA 1177
Qy 901 ATGACTCTGGAAGACAAAATTTGGATTGTGACAGAGGAGCTGCTCATATAATTAATGCC 960
Db 1178 ATGACTCTGGAAGACAAAATTTGGATTGTGACAGAGGAGCTGCTCATATAATTAATGCC 1237
Qy 961 ATGCAAAAAGGATCCACCTGACAGCTCTGAGGCGACTGTGAGAGTACAGAAAATGACG 1020
Db 1238 ATGCAAAAAGGATCCACCTGACAGCTCTGAGGCGACTGTGAGAGTACAGAAAATGACG 1297
Qy 1021 CCAAGAGAAAGGATCGGCTGAGAAAGCTCCAGGCGCTGATGAGAAAGCCAGGAAGCTG 1080
Db 1298 CCAAGAGAAAGGATCGGCTGAGAAAGCTCCAGGCGCTGATGAGAAAGCCAGGAAGCTG 1357
Qy 1081 AAAAGATTGTGGAAGAAAATATGAGAAAATAGCAAGAAATGCAAGAAATGAGATCT 1140
Db 1358 AAAAGATTGTGGAAGAAAATATGAGAAAATAGCAAGAAATGCAAGAAATGAGATCT 1417
Qy 1141 CGGAACCTTTACAGCAGCTGAGTGTGATGTACTCCATGAAAAACACATTTAAAGGAATG 1200
Db 1418 CGGAACCTTTACAGCAGCTGAGTGTGATGTACTCCATGAAAAACACATTTAAAGGAATG 1477
Qy 1201 GAGAAAAGGAGGAGCAACCTGAGGAGAGACTTCTTGTTCACCCAGGAGGAGGATGAA 1260
Db 1478 GAGAAAAGGAGGAGCAACCTGAGGAGAGACTTCTTGTTCACCCAGGAGGAGGATGAA 1537
Qy 1261 GAGAGGTGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
```

Db 1538 GAGAGTGGCAAGCGAGGAGGAATCTGATGAACCAACTTTAGAGAACCTGGCTGAG 1597

QY 1321 TCTCAGCCTATTCCTTCCATGGACCTCCAGAACTTGAATCAATTTAGAGGATGCCACA 1380

Db 1598 TCTCAGCCTATTCCTTCCATGGACCTCCAGAACTTGAATCAATTTAGAGGATGCCACA 1657

QY 1381 TCTCAGCCTTGGATACCATG 1399

Db 1658 TCTCAGCCTTGGATACCATG 1676

RESULT 5  
HSM804785

LOCUS HSM804785 2768 bp mRNA linear PRI 13-MAY-2003

DEFINITION Homo sapiens mRNA; cDNA DKF2p686L0617 (from clone DKF2p686L0617).

ACCESSION AL833472

VERSION AL833472.1 GI:21734115

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Bahr, A., Lauber, J., Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.

TITLE Direct Submission

JOURNAL Submitted (13-MAY-2003) MIPS, Ingolstaedter Landstr. 1, D-85764 Neuberg, GERMANY

COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing consortium of the German Genome Project.

This clone (DKF2p686L0617) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at <http://mips.gsf.de/proj/cDNA/>.

FEATURES

source

1..2768

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="DKF2p686L0617"

/tissue\_type="human cervix"

/clone\_lib="686 (synonym: hicc3). Vector pSport1\_Sfi; host DH10B; sites SfiIA + SfiIB"

/dev\_stage="adult"

polyA\_signal 2726..2731

polyA\_site 2753

ORIGIN

Query Match 96.5%; Score 1399; DB 9; Length 2768;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTGAATTCACAGACGACCTAAGTCTGTGAGTGAATCAACGCCATTTCCACTTAT 60

Db 225 ATGCTGAATTCACAGACGACCTAAGTCTGTGAGTGAATCAACGCCATTTCCACTTAT 284

QY 61 CCAAGACCTTGTATTCGAAGAAGATACGTGCTTCAACAAAACTTGGCAGTGGAAATTT 120

Db 285 CCAAGACCTTGTATTCGAAGAAGATACGTGCTTCAACAAAACTTGGCAGTGGAAATTT 344

QY 121 GGAAGTCTATCTGTTTTCAGACAAAGAGCCAAACGAGGAGAGGAATTAAGTACTT 180

Db 345 GGAAGTCTATCTGTTTTCAGACAAAGAGCCAAACGAGGAGAGGAATTAAGTACTT 404

QY 181 AAGGAATATCTGTTGGAGAACTAATCAATGAACTGTACAGGCCAATTTGGAAGCC 240

Db 405 AAGGAATATCTGTTGGAGAACTAATCAATGAACTGTACAGGCCAATTTGGAAGCC 464

QY 241 CAATCTCTTCCAGTGGACCAACCCAGCCATGTGCAAGTTCATGCAAGTTTGTGGAG 300

Db 465 CAATCTCTTCCAAAGCTGGACCAACCAGCCATTTGTCAAGTTCCATCAAGTTTGTGGAG 524

QY 301 CAAGATAAATTTCTGCATTATCAAGAGTACTGTGAGGGCCGAGATCTGGACGATAAAAT 360

Db 525 CAAGATAAATTTCTGCATTATCAAGAGTACTGTGAGGGCCGAGATCTGGACGATAAAAT 584

QY 361 CAGGAATATAAACAAGCTGGAAAAATCTTTCCAGAAATCAAAATTAATAGAAATGTTTATC 420

Db 585 CAGGAATATAAACAAGCTGGAAAAATCTTTCCAGAAATCAAAATTAATAGAAATGTTTATC 644

QY 421 CAGCTGCTGCTGGAGTTGACTACATGCTAGAGGAGGATATCTTCATCGAGACTTTAAAG 480

Db 645 CAGCTGCTGCTGGAGTTGACTACATGCTAGAGGAGGATATCTTCATCGAGACTTTAAAG 704

QY 481 TCAAGAAGATGATTTCTGAAAAATATCTCTTAAATTTGGAGATTTTGGAGTTTCTCGA 540

Db 705 TCAAGAAGATGATTTCTGAAAAATATCTCTTAAATTTGGAGATTTTGGAGTTTCTCGA 764

QY 541 CTTCTAATGGGATCCTGTGACCTGGCCACAACTTTAACTGGAACTCCCATTTATAGAT 600

Db 765 CTTCTAATGGGATCCTGTGACCTGGCCACAACTTTAACTGGAACTCCCATTTATAGAT 824

QY 601 CTTGAGGCTCTGAAAAACCAAGGCTATGACAAAGTGGACATCTGGTCACTGGGATGC 660

Db 825 CTTGAGGCTCTGAAAAACCAAGGCTATGACAAAGTGGACATCTGGTCACTGGGATGC 884

QY 661 ATTTTGTATGAGATGTCTGCTCATGATCATGATTCGCTGGCTCCAAATTTCTTATCCAT 720

Db 885 ATTTTGTATGAGATGTCTGCTCATGATCATGATTCGCTGGCTCCAAATTTCTTATCCAT 944

QY 721 GTTTTAAAAATTTGTTGAAGTGACACACCTTCTCTCCCTGAGAGATATCCAAAAGAACTA 780

Db 945 GTTTTAAAAATTTGTTGAAGTGACACACCTTCTCTCCCTGAGAGATATCCAAAAGAACTA 1004

QY 781 AATGCCATCATGAAAGCATGTTTGAACAGAACTCTTCAATTAAGACCATCTGCTATCGAA 840

Db 1005 AATGCCATCATGAAAGCATGTTTGAACAGAACTCTTCAATTAAGACCATCTGCTATCGAA 1064

QY 841 ATTTTAAAAATCCCTTACCTTGTATGAGCAGCTACAGAACCTAATGTGTAGATATTAGAA 900

Db 1065 ATTTTAAAAATCCCTTACCTTGTATGAGCAGCTACAGAACCTAATGTGTAGATATTAGAA 1124

QY 901 ATGACTCTGGAAGACAAAAATTTGGATTTGTGAGAGGAGCTCTCATATAATTAATGCC 960

Db 1125 ATGACTCTGGAAGACAAAAATTTGGATTTGTGAGAGGAGCTCTCATATAATTAATGCC 1184

QY 961 ATGCAAAAAAGGATCCACCTGACAGCTCTGAGGGCAGCTGTGAGAAATGAGGATGAGC 1020

Db 1185 ATGCAAAAAAGGATCCACCTGACAGCTCTGAGGGCAGCTGTGAGAAATGAGGATGAGC 1244

QY 1021 CCAAGAGAAAGGATGGCGCTGAGGAAGCTTCAGCGCGCTGTATGAGAAAGCCAGGAAGCTG 1080

Db 1245 CCAAGAGAAAGGATGGCGCTGAGGAAGCTTCAGCGCGCTGTATGAGAAAGCCAGGAAGCTG 1304

QY 1081 AAAAAAGTTCTGGAAGAAAAATATGAAGAAAAATAGCAAAACGAATCAAGAAATTCAGATCT 1140

Db 1305 AAAAAAGTTCTGGAAGAAAAATATGAAGAAAAATAGCAAAACGAATCAAGAAATTCAGATCT 1364

QY 1141 CGGAACCTTTCAGCAGCTGAGTGTGATGATCTCCATGAAAAACACATTTAAAGGAATG 1200

Db 1365 CGGAACCTTTCAGCAGCTGAGTGTGATGATCTCCATGAAAAACACATTTAAAGGAATG 1424

QY 1201 GAAGAAAGGAGAGCAACCTGAGGGAAGACTTTCTTTGTTTCACTCCAGACGAGGATGAA 1260

Db 1425 GAAGAAAGGAGAGCAACCTGAGGGAAGACTTTCTTTGTTTCACTCCAGACGAGGATGAA 1484

QY 1261 GAGAGTGGCAAGCGAGGGAAGGAAATCTGATGAACCAACTTTAGAGAACCTGCTGAG 1320

Db 1485 GAGAGTGGCAAGCGAGGGAAGGAAATCTGATGAACCAACTTTAGAGAACCTGCTGAG 1544

QY 1321 TCTCAGCCTATTCCTTCCATGGACCTCCAGACCTTGAATCAATTTAGAGGATGCCACA 1380

Db 1545 TCTCAGCCTATTCCTTCCATGGACCTCCAGACCTTGAATCAATTTAGAGGATGCCACA 1604

QY	1381	TTTGACCTTGGATACCATG	1399	QY	121	GGAACTGCTATCTGTTTCAGACAAGAAAGCCAAAGAGGAGGAAATTAAGGTACTT	180
Db	1605	TTTGACCTTGGATACCATG	1623	Db	414	GGAACTGCTATCTGTTTCAGACAAGAAAGCCAAAGAGGAGGAAATTAAGGTACTT	473
RESULT 6				QY	181	AAGGAAATATCTGTTGAGAACTAAATCCAAATGAAACTGTACAGGCCAATTTGGAAGCC	240
AB071996		AB071996	2939 bp mRNA linear PRI 17-OCT-2002	Db	474	AAGGAAATATCTGTTGAGAACTAAATCCAAATGAAACTGTACAGGCCAATTTGGAAGCC	533
LOCUS				QY	241	CAACTCTCTCCAGCTGGACCCAGCCATTTGCAAGTTCCATGCAAGTTTGTGTGAG	300
DEFINITION				Db	534	CAACTCTCTCCAGCTGGACCCAGCCATTTGCAAGTTTCCATGCAAGTTTGTGTGAG	593
ACCESSION				QY	301	CAAGATAATTTTCTGATTATCACGGAGTACTGTGAGGCGCCGAGATCTTGGACGATAAAAT	360
VERSION				Db	594	CAAGATAATTTTCTGATTATCACGGAGTACTGTGAGGCGCCGAGATCTTGGACGATAAAAT	653
KEYWORDS				QY	361	CAGGAATATAAACAAGCTGGAAAATCTTTCCAGAAAATCAATAATAGATGTTTTATC	420
SOURCE				Db	654	CAGGAATATAAACAAGCTGGAAAATCTTTCCAGAAAATCAATAATAGATGTTTTATC	713
ORGANISM				QY	421	CAGCTGCTGCTGGAGTTTGACTACATCATGACGAGGAGGATACCTTCAATCAGAGACTTAAAG	480
REFERENCE				Db	714	CAGCTGCTGCTGGAGTTTGACTACATCATGACGAGGAGGATACCTTCAATCAGAGACTTAAAG	773
AUTHORS				QY	481	TCAAAGAATGATATTTCTGAAAATTAATCTCTTTAAATTTGGAGATTTTGGAGTTTCTCGA	540
TITLE				Db	774	TCAAAGAATGATATTTCTGAAAATTAATCTCTTTAAATTTGGAGATTTTGGAGTTTCTCGA	833
JOURNAL				QY	541	CTTCTAATGGGATCTCTGTGACCTGGCCACAACTTTAACTGGAACCTCCCAATTTATATGAT	600
MEDLINE				Db	834	CTTCTAATGGGATCTCTGTGACCTGGCCACAACTTTAACTGGAACCTCCCAATTTATATGAT	893
PUBMED				QY	601	CCTGAGGCTCTGAAACACCAAGGCTATGACACAAAGTCGACATCTGGTCACTGGCATGC	660
FEATURES				Db	894	CCTGAGGCTCTGAAACACCAAGGCTATGACACAAAGTCGACATCTGGTCACTGGCATGC	953
source				QY	661	ATTTTGTATGAGATGCTGCGATGAATCATGCTTCGCTGGCTCCCAATTTCTTATCCATT	720
gene				Db	954	ATTTTGTATGAGATGCTGCGATGAATCATGCTTCGCTGGCTCCCAATTTCTTATCCATT	1013
CDS				QY	721	GTTTTAAAAATTTGTAAGGTGACACACCTTTCTCTCCCTGAGAGATATCAAAAGAACTA	780
				Db	1014	GTTTTAAAAATTTGTAAGGTGACACACCTTTCTCTCCCTGAGAGATATCAAAAGAACTA	1073
				QY	781	AATGCCATCATGAAAGCATGTTGAACAAGATCTTCAATTAAGACCATCTGCTATCGAA	840
				Db	1074	AATGCCATCATGAAAGCATGTTGAACAAGATCTTCAATTAAGACCATCTGCTATCGAA	1133
				QY	841	ATTTTAAAAATCCCTTACCTTGTAGCAGCTACAGAACCTTAATGCTGTAGATATTCAGAA	900
				Db	1134	ATTTTAAAAATCCCTTACCTTGTAGCAGCTACAGAACCTTAATGCTGTAGATATTCAGAA	1193
				QY	901	ATGACTCTGGAAGACAAAATTTGGATTGTCAGAGGAGGCTGCTATATAATTAATGCC	960
				Db	1194	ATGACTCTGGAAGACAAAATTTGGATTGTCAGAGGAGGCTGCTATATAATTAATGCC	1253
				QY	961	ATGCAAAAAAGGATCCACCTGCGACCTGTGAGGSCACTGTGCAAGGTACAGAAAATCAGC	1020
				Db	1254	ATGCAAAAAAGGATCCACCTGCGACCTGTGAGGSCACTGTGCAAGGTACAGAAAATCAGC	1313
				QY	1021	CCAAGAGAAGGATTCGGCTGAGGAAGCTCCAGCGGCTGATGAGAAGCCAGGAAGCTG	1080
				Db	1314	CCAAGAGAAGGATTCGGCTGAGGAAGCTCCAGCGGCTGATGAGAAGCCAGGAAGCTG	1373
				QY	1081	AAAAAGATTGGGAAGAAAATATGAAGAAAATAGCAAAAGTGAAGTGAAGTGAAGTCT	1140
				Db	1374	AAAAAGATTGGGAAGAAAATATGAAGAAAATAGCAAAAGTGAAGTGAAGTGAAGTCT	1433
				QY	1141	CGGAACCTTTTCAAGCAGCTGAGTGTGATGCTTCCATCAAAAAACACATTTTAAAGGAATG	1200
				Db	1434	CGGAACCTTTTCAAGCAGCTGAGTGTGATGCTTCCATCAAAAAACACATTTTAAAGGAATG	1493

ORIGIN

Query Match	96.5%	Score 1399;	DB 9;	Length 2939;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 1399;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
QY	1	ATGCTGAAATTCAGAGGAGCAGTAAAGTGTGTGAGTGAATCAACAGCCATTTCCACTTAT	60	
Db	294	ATGCTGAAATTCAGAGGAGCAGTAAAGTGTGTGAGTGAATCAACAGCCATTTCCACTTAT	353	
QY	61	CCAAAGACCTTGATTCAGAGAGATACGCTTCAACAAAACCTTGGCAGTGAAGTTT	120	
Db	354	CCAAAGACCTTGATTCAGAGAGATACGCTTCAACAAAACCTTGGCAGTGAAGTTT	413	



Db	955	ATTTTATAGAGATGCTGTCATGATCATGATTCGCTGGCTCCAAATTTCTATCCAT	1014
QY	721	GTTTTAAATTTTGAAGTGACACACCTTCTCTCCCTGAGAGATATCCAAAAGAACTA	780
Db	1015	GTTTTAAATTTTGAAGTGACACACCTTCTCTCCCTGAGAGATATCCAAAAGAACTA	1074
QY	781	AATGCCATCATGGAAGACATGTTGAACAAGATCCCTTCAATTAAGACCATCTGCTATCGAA	840
Db	1075	AATACATCATGGAAGACATGTTGAACAAGATCCCTTCAATTAAGACCATCTGCTATCGAA	1134
QY	841	ATTTTAAATTCCTTACCTTGATGAGAGCTACAGAACTAATGTTAGATATTCAGAA	900
Db	1135	ATTTTAAAGATCCCTTACATTTGATGAGAGCTACAGCACTAATGTTAGATATTCAGAA	1194
QY	901	ATGACTCTGGAACACAAAATTTGGATTTCTGAGAAGGAGCTGCTCATATTAATTCGCC	960
Db	1195	ATGACTCTGGAACACAAAATTTGGATTTCTGAGAAGGAGCTGCTCATATTAATTCGCC	1254
QY	961	ATGCAAAAAGGATCCACCTGACAGCTCTGAGGGCCTCTGCAAGCTACAGAAAATGAGC	1020
Db	1255	ATGCAAAAAGGATCCACCTGACAGCTCTGAGGGCCTCTGCAAGCTACAGAAAATGAGC	1314
QY	1021	CCNAGAGAAGGATGCGCTGAGGAAGCTCCAGGCGCTGATGAGAAACCCAGGAGCTG	1080
Db	1315	CCNAGAGAAGGATGCGCTGAGGAAGCTCCAGGCGCTGATGAGGAAGCCAGGAGCTG	1374
QY	1081	AAAAAATTTGGAACAAAAATATGAAGAAATAGCAAAAGATGCAAGATTTGAGATCT	1140
Db	1375	AAAAAATTTGGAAGAAAAATATGAAGAAATAGCAAAAGATGCAAGATTTGAGATCT	1434
QY	1141	CGAACTTTGAGAGCTGAGTGTGATCTGATCTCCATGAAAAACACATTTAAAGAAATG	1200
Db	1435	CGAACTTTGAGAGCTGAGTGTGATCTGATCTCCATGAAAAACACATTTAAAGAAATG	1494
QY	1201	GAAGAAAGGAGGAGCAACTGAGGAGCACTTCTGTTTCCACCCAGGAGGAGGATGAA	1260
Db	1495	GAAGAAAGGAGGAGCAACTGAGGAGCACTTCTGTTTCCACCCAGGAGGAGGATGAA	1554
QY	1261	GAGAGTGCAAGGAGGAGGAGGATCTGATGACCAACTTTAGAACCTGCTGCTGAG	1320
Db	1555	GAGAGTGCAAGGAGGAGGAGGATTTGATGACCGACTTTAGAACCTGCTGCTGAG	1614
QY	1321	TCTGACCTATTCCTTCCATGAGCTCCAGCACTGATGATGATGATGAGAGATGCCACA	1380
Db	1615	CCTGACCTATTCCTTCCATGAGCTCCAGCACTGATGATGATGATGAGAGATGCCACA	1674
QY	1381	TCTGACCTTGGATACCATG 1399	
Db	1675	TCTGACCTTGGATATCATG 1693	

RESULT 8  
AK027148  
LOCUS AK027148 2049 bp mRNA linear PRI 12-SEP-2003  
DEFINITION Homo sapiens cDNA: FLJ23495 fis, clone LNG02228.  
ACCESSION AK027148  
VERSION AK027148.1 GI:10440201  
KEYWORDS oligo capping; fls (full insert sequence).  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Kawakami, T., Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hirao, M., Ohmori, Y., Ota, T., Suzuki, Y., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.  
NEDO human cDNA sequencing project  
Unpublished  
TITLE NEDO human cDNA sequencing project  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 2049)  
AUTHORS Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Isogai, T.,

Shibahara, T., Tanaka, T. and Nakamura, Y.  
Direct Submission  
Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail: flicdha@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)  
NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).  
Location/Qualifiers  
1..2049  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="LNG02228"  
/tissue\_type="human lung"  
/clone\_lib="LNG"  
/note="cloning vector pME18SFL3"  
66..1457  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="BAB:5672.1"  
/db\_xref="GI:10440202"  
/translation="MGSCDLATLTGTPTHTSPALXHQGYDTKSDIWSLACILYEMC  
CMNFAFGSNFLSIVLKIVGDTPLSPRPKELNAMESLNKNSPRLSRAELIKI  
PYDDEQLNLMCYSEMTEDKNDLCOKEAAHINAMQKRIHQITLSEVQKQTPR  
EVRRLKQADAEKARKLKIVKEEYENSRKMOELSRNFQOLSVDVLHEKTHKGM  
EEKEQPEGLSCSPDEDEERWQREESDEPTLENLPSQIPMDLHELSEIVED  
ATSLGVHEIPEDPLVAEYVADFSDYSCVDEDEEEIALERPEKIRNEGSOAYR  
TNODSDIEALALSCLENVLGTSIDTKITTMEDMSPPPIENSWARTKMKEMRES  
AMQKLGEVEVEVYNYLKRARHQAASEARECLEKVKVVPQASDCFEVDQLLYEEQLL  
ITMKEPTLNHLL"  
ORIGIN  
Query Match 63.4%; Score 918; DB 9; Length 2049;  
Best Local Similarity 100.0%; Pred. No. 3e-222;  
Matches 918; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 482 CAAGAATGATTTCTGAAAAATAATCTCTTAAAAATGGAGATTTGGAGTTTCGCAC 541  
Db 1 CAAGAATGATTTCTGAAAAATAATCTCTTAAAAATGGAGATTTGGAGTTTCGCAC 60  
QY 542 TTCTAATGGGATCTGTGACCTGGCCACAACTTTAACTGGAATCCCATTTATATGATG 601  
Db 61 TTCTAATGGGATCTGTGACCTGGCCACAACTTTAACTGGAATCCCATTTATATGATG 120  
QY 602 CTGAGGCTCTGAAACACCAAGGCTATGACACAAAGTCGGACATCTGTCTACTGGCATGCA 661  
Db 121 CTGAGGCTCTGAAACACCAAGGCTATGACACAAAGTCGGACATCTGTCTACTGGCATGCA 180  
QY 562 TTTTGTATGAGATGCTGCGATGATCATGATCGCTGGCTGCTCAATTTCTTATCCATTG 721  
Db 181 TTTTGTATGAGATGCTGCGATGATCATGATCGCTGGCTGCTCAATTTCTTATCCATTG 240  
QY 722 TTTTAAAAATTTGTTGAAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAAGAACTAA 781  
Db 241 TTTTAAAAATTTGTTGAAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAAGAACTAA 300  
QY 782 ATGCCATCATGGAAGCATGTTGAAAGAATCTCTCAATTAAGACCATCTGCTATCGAAA 841  
Db 301 ATGCCATCATGGAAGCATGTTGAAAGAATCTCTCAATTAAGACCATCTGCTATCGAAA 360  
QY 842 TTTTAAAAATTCCTTACCTTGATGAGCAGCTACAGAACCTTAATGTTGATGATATTCAGAAA 901  
Db 361 TTTTAAAAATTCCTTACCTTGATGAGCAGCTACAGAACCTTAATGTTGATGATATTCAGAAA 420  
QY 902 TGACTCTGGAAGCAAAAATTTGATGTTGAGAGGAGCTGCTCATATTAATTAATGCCA 961







QY 601 CCTGAGGCTCTGAACACCAAGGCTATGACACAAAGTCGACATCTGTCATCGCATGC 660  
Db 766 CTTGAGGCTTTGAACACCAAGGCTATGACACAAAGTCGACATCTGTCATCGCATGC 825  
QY 661 ATTTTGTATGAGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATG 720  
Db 826 ATTTTATATGAGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATG 885  
QY 721 GTTTTAAATTTGTTGAAGGTGACACACCTTCTCTGCTGAGAGATATCCAAAGACTA 780  
Db 886 GTTTTAAATTTGTTGAAGGTGACACACCTTCTCTGCTGAGAGATATCCAAAGACTA 945  
QY 781 AATGCCATCATGGAAGAGCTGTTGAACAAGATCCCTTCAATTAAGACCACTGCTATCGAA 840  
Db 946 AATGCCATCATGGAAGAGCTGTTGAACAAGATCCCTTCAATTAAGACCACTGCTATCGAA 1005  
QY 841 ATTTTAAATTTCCCTTACCTTGTATGATGATGATGATGATGATGATGATGATGATG 900  
Db 1006 ATTTTAAAGATCCCTTACCTTGTATGATGATGATGATGATGATGATGATGATGATG 1065  
QY 901 ATGACTCTGGAAGACAAAATTTGGATTTGTGCAAGGAGGCTGCTCATATAATTAATGCC 960  
Db 1066 ATGACTCTGGAAGACAAAATTTGGATTTGTGCAAGGAGGCTGCTCATATAATTAATGCC 1125  
QY 961 ATGCAAAAAGGATCCACTGCGAGACTCTGAGGCACTGTGCAAGATGACAGAAATGACG 1020  
Db 1126 ATGCAAAAAGGATCCACTGCGAGACTCTGAGGCACTGTGCAAGATGACAGAAATGACG 1185  
QY 1021 CCAAGAGAAGGATGCGCTGAGGAAGCTCCAGGCGCTGATGAGAAAGCCAGGAGCTG 1080  
Db 1186 CCAAGAGAAGGATGCGCTGAGGAAGCTCCAGGCGCTGATGAGAGAGCCAGGAGCTG 1245  
QY 1081 AAAAGATTGTGGAAGAAAATATGAAGAAAATAGCAAAAGATGCAAGAAATGAGATCT 1140  
Db 1246 AAAAGATTGTGGAAGAAAATATGAAGAAAATAGCAAAAGATGCAAGAAATGAGATCT 1305  
QY 1141 CGGAATTTTACGAGCTGATGTTGATGATGATGATGATGATGATGATGATGATGATG 1200  
Db 1306 CGGAATTTTACGAGCTGATGTTGATGATGATGATGATGATGATGATGATGATGATG 1365  
QY 1201 GAAGAAAAGGAGAGCAACTGAGGGAAGACTTTCTTTTACCCCGAGGACAGGATGAA 1260  
Db 1366 GAAGAAAAGGAGAGCAACTGAGGGAAGACTTTCTTTTACCCCGAGGACAGGATGAA 1425  
QY 1261 GAGAGTGGCAGGAGGAGGAGGATCTGATGATGATGATGATGATGATGATGATGATG 1320  
Db 1426 GAGAGTGGCAGGAGGAGGAGGATTTGATGATGATGATGATGATGATGATGATGATG 1485  
QY 1321 TCTCAGCCTATTCTTCCATGAGCTCCAGCACTTGAATCAATTTGATGAGGATGCCACA 1380  
Db 1486 CTTGACCTTTCTTCCATGAGCTCCAGCACTTGAATCAATTTGATGAGGATGCCACA 1545  
QY 1381 TCTGACCTTTGATACCATG 1399  
Db 1546 TCTGACCTTTGATATCATG 1564

RESULT 10  
AK123021  
LOCUS  
DEFINITION Homo sapiens cDNA FLJ41025 fis, clone BRACE2008594, weakly similar  
to Human protein serine/threonine kinase stk2 mRNA.  
ACCESSION AK123021  
VERSION AK123021.1 GI:34528469  
KEYWORDS Oligo capping; fis (full insert sequence).  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE  
AUTHORS Tashiro,H., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S.,  
Fukuzumi,Y., Fujimori,Y., Komiyama,M., Sugiyama,T., Irie,R.,  
Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,

Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,  
Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,  
Wagatsuma,M., Murakawa,K., Kurehori,K., Takahashi-Fujii,A.,  
Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S.,  
Nagahara,K., Masuho,Y., Nagai,K. and Isogai,T.  
NEBO human cDNA sequencing project  
Unpublished  
2 (bases 1 to 2189)  
Isogai,T. and Yamamoto,J.  
Direct Submission  
Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7  
Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan  
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)  
NEBO human cDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan; cDNA full insert sequencing:  
Research Association for Biotechnology (RAB); cDNA library  
construction: Helix Research Institute (HRI) (supported by Japan  
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,  
HRI, and Biotechnology Center, National Institute of Technology and  
Evaluation; clone selection for full insert sequencing: HRI and  
RAB; annotation: HRI and RAB.  
FEATURES  
source  
1..2189  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="BRACE2008594"  
/tissue\_type="cerebellum"  
/clone\_lib="BRACE2"  
/note="Cloning vector: pME18SFL3"  
Query Match 45.0%; Score 651.8; DB 9; Length 2189;  
Best Local Similarity 99.7%; Pred. No. 1.4e-154;  
Matches 653; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ATGCTGAAATCCAAAGAGGCGAGCTAAGTGTGTGAGTGGATCAACAGCCATTTCCACTTAT 60  
Db 1221 ATGCTGAAATCCAAAGAGGCGAGCTAAGTGTGTGAGTGGATCAACAGCCATTTCCACTTAT 1280  
QY 61 CCAAGACCTTGTATTGCAAGAAGATAGTGTCTCAACAAAACCTTGGCAGTGGAGTTT 120  
Db 1281 CCAAGACCTTGTATTGCAAGAAGATAGTGTCTCAACAAAACCTTGGCAGTGGAGTTT 1340  
QY 121 GGAATCTCTATCTGTTTTCAGCAAGAAAGCCAAACGAGGAGGAGGAATTAAGGTACTT 180  
Db 1341 GGAATCTCTATCTGTTTTCAGCAAGAAAGCCAAACGAGGAGGAGGAATTAAGGTACTT 1400  
QY 181 AAGGAATATCTGTGTGAGAACTAAATCCAAATGAACTGTACAGGCCAATTTGGAGCC 240  
Db 1401 AAGGAATATCTGTGTGAGAACTAAATCCAAATGAACTGTACAGGCCAATTTGGAGCC 1460  
QY 241 CAACTCTCTCCAGCTGGACCCAGCCATTTGTCAAGTTCCATGCAAGTTTGTGGAG 300  
Db 1461 CAACTCTCTCCAGCTGGACCCAGCCATTTGTCAAGTTCCATGCAAGTTTGTGGAG 1520  
QY 301 CAAGATAATTTCTGCATTATCAGGAGTACTGTGAGGCCGAGATCTGGACGATAAAATT 360  
Db 1521 CAAGATAATTTCTGCATTATCAGGAGTACTGTGAGGCCGAGATCTGGACGATAAAATT 1580  
QY 361 CAGGAATATAACCAAGCTGGAATAATCTTTCCAGAAAATCAAATAATAGAAATGTTTATC 420  
Db 1581 CAGGAATATAACCAAGCTGGAATAATCTTTCCAGAAAATCAAATAATAGAAATGTTTATC 1640  
QY 421 CAGCTGTCTGGAGTTGACTCATGATGAGAGGAGGATATCTTCATCAGAGCTTAAAG 480  
Db 1641 CAGCTGTCTGGAGTTGACTCATGATGAGAGGAGGATATCTTCATCAGAGCTTAAAG 1700  
QY 481 TCAAGAAATGTAATTTCTGAAAATAATCTCTTAAATTCGAGATTTGGAGTTTCTCGA 540  
Db 1701 TCAAGAAATGTAATTTCTGAAAATAATCTCTTAAATTCGAGATTTGGAGTTTCTCGA 1760  
QY 541 CTTCTAATGGGATCCTGTGACCTGGCCACCAACTTTTAACCTGGAACCTCCCATTTATGAGT 600

```

Db      1761  CTTCTAATGGATCTGTGACCTGGCCACAACTTTAACTGGAACTCCCATATATAGT 1820
QY      601  CTTGAGGCTCTGAACACCAAGCTATGACACAAAGTCGGACATCTGTCACGTGG 655
Db      1821  CTTGAGGCTCTGAACACCAAGCTATGACACAAAGTCGGACATCTGTCAGTGG 1875

RESULT 11
AC025148/c
LOCUS   AC025148.3 GI:8077076          154330 bp      DNA      linear      HTG 24-AUG-2002
DEFINITION Homo sapiens chromosome 3 clone RP11-265F19 map 3, WORKING DRAFT
SEQUENCE, 11 unordered pieces.
ACCESSION AC025148
VERSION   AC025148.3
KEYWORDS  HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE   Homo sapiens
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 154330)
          Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
          Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
          Boguslavskiy, L., Boukhalter, B., Brown, A., Burkett, G.,
          Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
          Collamore, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S.,
          Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,
          Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
          Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
          Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
          Klein, J., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J.,
          Levine, R., Liu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
          McElrath, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,
          Meldrim, J., Meneus, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J.,
          Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
          O'Neil, D., Olivier, T.M., Oliver, J., Peterson, K., Pierre, N.,
          Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
          Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
          Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
          Tesfayev, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
          Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
          Young, G., Zainoun, J., Zimmer, A. and Zody, M.
          Direct Submission
          Submitted (05-MAR-2000) Whitehead Institute/MIT Center for Genome
          Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 154330)
          Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
          Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
          Boguslavskiy, L., Boukhalter, B., Brown, A., Burkett, G.,
          Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,
          Collamore, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S.,
          Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,
          Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
          Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
          Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
          Klein, J., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J.,
          Levine, R., Liu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
          McElrath, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,
          Meldrim, J., Meneus, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J.,
          Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
          O'Neil, D., Olivier, T.M., Oliver, J., Peterson, K., Pierre, N.,
          Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
          Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
          Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
          Tesfayev, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
          Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
          Young, G., Zainoun, J., Zimmer, A. and Zody, M.
          Direct Submission
          Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
  
```

```

Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 25, 2000 this sequence version replaced gi:7272269.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L7980
Center clone name: 265_F_19
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 149112 bases at least Q40
Consensus quality: 151109 bases at least Q30
Consensus quality: 152128 bases at least Q20
Insert size: 150000; agarose-fp
Insert size: 153330; sum-of-contigs
Quality coverage: 5.6 in Q20 bases; agarose-fp
Quality coverage: 5.5 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1      1829: contig of 1829 bp in length
*      1929: gap of 100 bp
*      5551: contig of 3622 bp in length
*      5651: gap of 100 bp
*      5652: contig of 6148 bp in length
*      11899: gap of 100 bp
*      20726: contig of 8827 bp in length
*      20826: gap of 100 bp
*      33161: contig of 12335 bp in length
*      33261: gap of 100 bp
*      46620: contig of 13359 bp in length
*      46720: gap of 100 bp
*      63335: contig of 16615 bp in length
*      63336: gap of 100 bp
*      81256: contig of 17821 bp in length
*      81257: gap of 100 bp
*      81357: contig of 19169 bp in length
*      100526: gap of 100 bp
*      122826: contig of 22201 bp in length
*      122827: gap of 100 bp
*      154330: contig of 31404 bp in length.
*
* Location/Qualifiers
*       1..154330
*         /organism="Homo sapiens"
*         /mol_type="genomic DNA"
*         /db_xref="taxon:9606"
*         /chromosome="3"
*         /map="3"
*       /clone="RP11-265F19"
*       /clone_lib="RPC1-11 Human Male BAC"
*       1..1829
*         /note="assembly_fragment"
*       1930..5551
*         /note="assembly_fragment"
*       5652..11799
*         /note="assembly_fragment"
*       11900..20726
*         /note="assembly_fragment"
*       20827..33161
*         /note="assembly_fragment"
  
```

```

FEATURES
source
misc_feature
misc_feature
misc_feature
misc_feature
misc_feature
  
```

```

misc_feature 33262..46620
/note="assembly_fragment"
misc_feature 46721..63335
/note="assembly_fragment"
misc_feature 63436..81256
/note="assembly_fragment"
clone_end:SP6
vector_side:right"
misc_feature 81357..100525
/note="assembly_fragment"
misc_feature 100626..122826
/note="assembly_fragment"
clone_end:TV
vector_side:left"
misc_feature 122927..154330
/note="assembly_fragment"
ORIGIN
Query Match 11.9%; Score 172.2; DB 2; Length 154330;
Best Local Similarity 98.3%; Pred. No. 1-5e-32;
Matches 174; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGCTGAATTCACAGGAGCAGTAAGTGTGTGAGTGATCAACAGCCATTTCCACTTAT 60
DB 137522 ATGCTGAATTCACAGGAGCAGTAAGTGTGTGAGTGATCAACAGCCATTTCCACTTAT 137463

QY 61 CCAAGACCTTGATTCACAGAGATACGCTTCAACAAAACCTGGCAGTGGAGTTT 120
DB 137462 CCAAGACCTTGATTCACAGAGATACGCTTCAACAAAACCTGGCAGTGGAGTTT 137403

QY 121 GGAATGCTATCTGTTTTCAGACAGAAAGCCAAACGAGGAGAGAAATTAAGGTA 177
DB 137402 GGAATGCTATCTGTTTTCAGACAGAAAGCCAAACGAGGAGAGAAATTAAGGTA 137346

RESULT 12
AC055733/c
LOCUS Homo sapiens 3 BAC RP11-39E3 (Roswell Park Cancer Institute Human
DEFINITION BAC Library) complete sequence.
ACCESSION AC055733
VERSION AC055733.16 GI:18463999
KEYWORDS Htg.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 193425)
Muzny,D.M., Adams,C., Aio-Oduola,B., Ali-oman,F.R., Allen,C.,
Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayete,M., Banks,T.,
Barbata,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carson,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Penn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorelli,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X.,
Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M., Holloway,C.,
Hollins,B., Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Louiseged,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,

```

```

Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D.,
Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N.,
Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N.,
Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L.,
Pickens,A., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M.,
Rojas,A., Rojebokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S.,
Scott,G., Shen,H., Shoshitari,N., Sisson,I., Sodargren,E.,
Sonaik,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A.,
Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C.,
Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L.,
Vera,V., Villalon,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S.,
Warren,R., Washington,C., Watlington,S., Williams,G.,
Williamson,A., Wlezyk,R., Woodden,S., Worley,K., Wu,C., Wu,Y.,
Wu,Y.F., Zhou,J., Zorrilla,S., Zorrilla,S., Zorrilla,S., Zorrilla,S.,
Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 193425)
Worley,K.C.
Direct Submission
Submitted (18-APR-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 193425)
Worley,K.C.
Direct Submission
Submitted (01-FEB-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 193425)
Worley,K.C.
Direct Submission
Submitted (28-MAR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 193425)
Worley,K.C.
Direct Submission
Submitted (30-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Feb 1, 2002 this sequence version replaced gi:17432818.
INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
gc-help@bcm.tmc.edu
CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.
ANNOTATION OF FEATURES:
STSS are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
EST and cDNA sequences. Genes demonstrate at least two exons
flanked by consensus splice sites that maintained sequence
continuity across the splice junctions. Sequences that are not
identical matches are annotated as similar.
SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads, with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as Low Coverage.

```

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

# FEATURES

QUALSTAT-REPORT.

Location/Qualifiers

1. 193425

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="3"

/clone="RP11-39E3"

2392. 2487

/rpt\_family="CATATA"

complement(3502..3636)

5830. 5874

/rpt\_family="L1MA10"

6418. 6709

/rpt\_family="AT-rich"

7114. 7164

/rpt\_family="AluSc"

7165. 7340

/rpt\_family="L2"

7341. 7400

/rpt\_family="MER5A"

7532. 7684

/rpt\_family="L2"

9312. 9592

/rpt\_family="MER5B"

/rpt\_family="AluY"

complement(11174..11303)

11554. 11736

/rpt\_family="MIR"

11737. 12033

/rpt\_family="MER102"

12034. 12101

/rpt\_family="AluX"

12322. 12366

/rpt\_family="MER102"

complement(13748..14040)

15145. 15443

/rpt\_family="AluX"

16404. 16477

/rpt\_family="AluY"

17769. 17891

/rpt\_family="L2"

complement(17817..17958)

complement(20059..20212)

20622. 20650

/rpt\_family="MIR"

20926. 21132

/rpt\_family="AT-rich"

21275. 21376

/rpt\_family="MER20"

21381. 21638

/rpt\_family="L1MB6"

21644. 21710

/rpt\_family="AluJo"

complement(22153..22307)

complement(22532..22708)

complement(22787..23078)

complement(23397..23463)

/rpt\_family="L1MB7"

/rpt\_family="L1MA10"

/rpt\_family="L1ME3"

/rpt\_family="L1M4"

/rpt\_family="MIR"

repeat\_region 24865..25001

/rpt\_family="MER5B"

repeat\_region complement(24884..25018)

/rpt\_family="MER5A"

repeat\_region complement(25637..25720)

/rpt\_family="L2"

repeat\_region complement(26002..26308)

/rpt\_family="AluSg"

repeat\_region complement(26475..26650)

/rpt\_family="MER3"

STS 27663..27812

/standard\_name="51335"

repeat\_region 28156..28354

/rpt\_family="L2"

repeat\_region complement(28308..28663)

/rpt\_family="L2"

repeat\_region 28665..28752

/rpt\_family="L2"

Query Match 11.98; Score 172.2; DB 9; Length 193425;

Best Local Similarity 98.3%; Pred. No. 1.6e-32;

Matches 174; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGCTGAATTCCTCAAGAGGAGGAGCTAAGTGTGTGAGTGGATCAACACGCAATTTCCACTTAT 60

DB 14289 ATGCTGAATTCCTCAAGAGGAGGAGCTAAGTGTGTGAGTGGATCAACACGCAATTTCCACTTAT 14230

QY 61 CCAAGACCTTGTATTCAGAGAGATAGTGTGTGAGTGGATCAACACGCAATTTCCACTTAT 120

DB 14229 CCAAGACCTTGTATTCAGAGAGATAGTGTGTGAGTGGATCAACACGCAATTTCCACTTAT 14170

QY 121 GGAATCTCTATCTGTTTCAGACAGAAAGCCAAACGAGGAGGAGAAATTAAGTA 177

DB 14169 GGAATCTCTATCTGTTTCAGACAGAAAGCCAAACGAGGAGGAGAAATTAAGTA 14113

RESULT 13

AC116424/c 60398 bp DNA linear PRI 28-JUN-2002

LOCUS AC116424 Homo sapiens 3 BAC RP11-265F19 (Roswell Park Cancer Institute Human

DEFINITION BAC library) complete sequence.

ACCESSION AC116424

VERSION AC116424.1 GI:19774299

KEYWORDS HTG.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 60398)

Muszy D.M., Adams C., Adio-Oduola B., Ali-oshan F.R., Allen C., Alsbrooks S.L., Amarutunge H.C., Are J.R., Ayele M., Banks T., Barbara J., Benton J., Bimaga K., Blankenburg K., Bonnin D., Bouck J., Bowie S., Brieva M., Brown E., Brown M., Bryant N.P., Buhay C., Burch P., Burkett C., Burrell K.L., Byrd N.C., Caron T.F., Carter M., Cavazos S.R., Chacko J., Chavez D., Chen G., Chen R., Chen Z., Chowdhry I., Christopoulos C., Cleveland C.D., Cox C., Coyle M.D., Dathorne S.R., David R., Davila M.L., Davis C., Davy-Carroll L., Dederich D.A., Delaney K.R., Delgado O., Denn A.L., Ding Y., Dinh H.H., Douthwaite K.J., Draper H., Dugan-Rocha S., Durbin K.J., Earnhart C., Edgar D., Edwards C.C., Elhaj C., Escotto M., Falls T., Ferraguto D., Flagg N., Ford J., Foster P., Frantz P., Gabisi A., Gao J., Garcia A., Garner T., Garza N., Gill R., Gorrell J.H., Guevara W., Gunaratne P., Hale S., Hamilton K., Harris C., Harris K., Hart M., Havlak P., Hawes A., He X., Hernandez J., Hernandez O., Hodgson A., Hogue M., Holloway C., Hollins B., Homs F., Howard S., Huber J., Hulyk S., Hume J., Jackson L.E., Jacobson B., Jia Y., Johnson R., Jolivet S., Joudah S., Karlsson E., Kelly S., Khan U., King L., Korvah J., Kovar C., Kratovic J., Kureshi A., Landry N., Leal B., Lewis L.C., Lewis L., Li J., Li Z., Lichtarge O., Lied C., Liu J., Liu W., Louis H., Lozano R., Lu X., Lucier A., Lucier R., Luna R., Ma J., Maheshwari M., Mapua P., Martin R., Martindale A.,

Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Mitchel,Z., Mitchell,T., Mohabbat,K., Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Reni,F., Rives,M., Rojias,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I., Sodergren,E., Sotaike,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,T., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Vallalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczky,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstein,G. and Gibbs,R.

Direct Submission  
Unpublished  
2 (bases 1 to 60398)  
Worley,K.C.

Direct Submission  
Submitted (28-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 60398)

Direct Submission  
Submitted (30-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
4 (bases 1 to 60398)  
Worley,K.C.

Direct Submission  
Submitted (28-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email [gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

#### ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDS, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

(Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:

<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

FEATURES	QUALSTAT-REPORT.	Location/Qualifiers
source	1..60398	/organism="Homo sapiens"
		/mol_type="Genomic DNA"
		/db_xref="taxon:9606"
		/chromosome="3"
		/clones="RP11-265F19"
misc_feature	complement(1..2005)	/note="overlaps bases 1..2005 of clone AC121332"
		/function="clone overlap"
repeat_region	13..1803	/rpt_family="L1P4"
repeat_region	complement(1804..2070)	/rpt_family="AluSg"
repeat_region	2071..3576	/rpt_family="L1P4"
repeat_region	3577..3807	/rpt_family="AluSg/x"
repeat_region	3808..3996	/rpt_family="L1P4"
repeat_region	complement(3976..4781)	/rpt_family="L1P"
STS	4444..4621	/standard_name="61535"
repeat_region	4783..5641	/rpt_family="L1P3"
repeat_region	5644..5680	/rpt_family=" (TTAAA)n"
repeat_region	5681..6309	/rpt_family="L1P16"
repeat_region	6433..7738	/rpt_family="MERS7-internal"
repeat_region	8070..8214	/rpt_family="MERS7-internal"
repeat_region	8241..8544	/rpt_family="MERS7-internal"
repeat_region	8680..8954	/rpt_family="AluSg"
repeat_region	9458..9582	/rpt_family="U6"
repeat_region	9577..10252	/rpt_family="MERS7-internal"
repeat_region	10284..10645	/rpt_family="MERS7B"
repeat_region	10999..11020	/rpt_family="AT-rich"
repeat_region	11322..11627	/rpt_family="AluJo"
repeat_region	12348..12369	/rpt_family=" (CAAAA)n"
STS	12449..12646	/standard_name="173989"
repeat_region	12851..12957	/rpt_family="AluSx"
repeat_region	complement(13041..13219)	/rpt_family="MERSA"
repeat_region	14844..15152	/rpt_family="AluSc"
repeat_region	15679..15779	/rpt_family="MIR"
repeat_region	15782..16643	/rpt_family="L1MC4"
repeat_region	16688..16826	/rpt_family="L2"
repeat_region	17461..17482	/rpt_family="AT-rich"
repeat_region	17487..17655	/rpt_family="MERSB"
repeat_region	complement(18331..18445)	/rpt_family="L2"
repeat_region	complement(19199..19389)	

```

/rpt_family="MER5A"
19396..19676
/rpt_family="LMC4"
complement(19733..19937)
/rpt_family="MER3"
complement(21247..21557)
/rpt_family="AluJo"
22046..22393
/standard_name="173331"
22175..22391
/standard_name="61389"
22252..22391
/standard_name="59650"
22284..22315
/rpt_family="(CA)n"
22316..22359
/rpt_family="(CA)n"
23450..23482
/rpt_family="AT_rich"

Query Match 11.5%; Score 167.2; DB 9; Length 60398;
Best Local Similarity 98.3%; Pred. No. 2.6e-31;
Matches 169; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 171 AAAGGTACTTAAGGAATATCTTGGAGAACTTAATCAATGAACTGTACAGGCCAA 230
Db 21962 AAAGGTACTTAAGGAATATCTTGGAGAACTTAATCAATGAACTGTACAGGCCAA 21903

Qy 231 TTGTGAAGCCCACTCTCTCCAAAGTGGACCAACCAGCCATGTCAAGTTCATGCAAG 290
Db 21902 TTGTGAAGCCCACTCTCTCCAAAGTGGACCAACCAGCCATGTCAAGTTCATGCAAG 21843

Qy 291 TTGTGTGGACGACGATAATTTCTGCATTATCAAGGAGTACTGTGAGGGCGGA 342
Db 21842 TTGTGTGGACGACGATAATTTCTGCATTATCAAGGAGTACTGTGAGGGCGGA 21791

RESULT 14
AC083908/c 111568 bp DNA linear PRI 28-JUN-2002
LOCUS Homo sapiens 3 BAC RP11-431K13 (Roswell Park Cancer Institute Human
DEFINITION BAC library) complete sequence.
ACCESSION AC083908
VERSION AC083908.19 GI:21263155
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaraturge,H.C., Are,J.R., Ayale,M., Banks,T.,
Barbata,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
DeLaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J.J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gerrill,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X.,
Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Holloway,C.,
Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
Joudan,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,

```

```

Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,
Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D.,
Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N.,
Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N.,
Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L.,
Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M.,
Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S.,
Scott,G., Shen,H., Shooshtari,N., Sisson,I., Sodergren,E.,
Sonaika,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A.,
Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C.,
Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,I.,
Vara,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S.,
Warren,R., Washington,C., Watlington,S., Williams,G.,
Williamson,A., Wlarczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y.,
Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstein,G. and
Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 111568)
Worley,K.C.
Direct Submission
Submitted (06-OCT-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 111568)
Worley,K.C.
Direct Submission
Submitted (21-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 111568)
Worley,K.C.
Direct Submission
Submitted (30-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 111568)
Worley,K.C.
Direct Submission
Submitted (28-JUN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 30, 2002 this sequence version replaced gi:20976448.
INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
gc-help@bcm.tmc.edu

```

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

#### ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the



Db 104773 TCCTGAGAGATATCCAAAGAACTAAATGCCATCATGGAAAG 104731

Search completed: October 2, 2004, 23:48:58  
Job time : 5798 secs

----- Project Information  
Center project name: 888  
Center clone name: RP11-290D10  
----- Summary Statistics  
Sequencing Vector: M13mp18; X2513; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Sequencing Vector: plasmid; plasmid\_accession; 0% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 171536 bases at least Q40  
Consensus quality: 173017 bases at least Q30  
Consensus quality: 173757 bases at least Q20  
Insert size: 16599; agarose-fp  
Insert size: 17542; sum-of-contigs  
Quality coverage: 11.9x in Q20 bases; agarose-fp  
Quality coverage: 11.5x in Q20 bases; sum-of-contigs.

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 6 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 1839: contig of 1839 bp in length  
\* 1840 1839: gap of unknown length  
\* 1840 3792: contig of 1853 bp in length  
\* 3793 3892: gap of unknown length  
\* 3893 6447: contig of 2555 bp in length  
\* 6448 6548: gap of unknown length  
\* 6548 8332: contig of 1785 bp in length  
\* 8333 8432: gap of unknown length  
\* 8433 25713: contig of 17281 bp in length  
\* 25714 25813: gap of unknown length  
\* 25814 175923: contig of 150110 bp in length.

## FEATURES

source  
1..175923  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="3"  
/clone="RP11-290D10"  
/clone\_lib="RPCI human BAC library 11"  
1..1839  
/note="assembly\_name:Contig36"  
1840..3792  
/note="assembly\_name:Contig41"  
3893..6447  
/note="assembly\_name:Contig44"  
6548..8332  
/note="assembly\_name:Contig46"  
8433..25713  
/note="assembly\_name:Contig47"  
clone\_end:77"  
25814..175923  
/note="assembly\_name:Contig48"  
clone\_end:SP6"

## ORIGIN

Query Match 10.6%; Score 153.4; DB 2; Length 175923;  
Best Local Similarity 96.3%; Pred. NO. 9.3e-28;  
Matches 157; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 635 AGTCGACATCTCGTCACATGCGATGATTTGATGAGATGCTGCGATCATGTCAT 694  
Db 104893 ACITTTGCTTCAGTCACTGGCATGTCATTTTGTATGAGATGCTGCGATCATGTCAT 104834  
QY TCCTGGCTCCATTTCTTATCCATTTGTTTAAATTTGTTGAAGTGACACACCTTCTC 754  
Db 104833 TCCTGGCTCCATTTCTTATCCATTTGTTTAAATTTGTTGAAGTGACACACCTTCTC 104774  
QY 755 TCCTGAGAGATATCCAAAGAACTAAATGCCATCATGGAAAG 797



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2004, 17:05:50 ; Search time 4039 Seconds

(without alignments)  
10713.122 Million cell updates/sec

Title: US-10-803-278-3

Perfect score: 1449

Sequence: 1 atgctgaattccaagagc.....actggaaaatgaaaaataa 1449

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_ges\_hum:\*  
18: em\_ges\_inv:\*  
19: em\_ges\_pln:\*  
20: em\_ges\_vrt:\*  
21: em\_ges\_fun:\*  
22: em\_ges\_mam:\*  
23: em\_ges\_mus:\*  
24: em\_ges\_pro:\*  
25: em\_ges\_rod:\*  
26: em\_ges\_pbg:\*  
27: em\_ges\_vrl:\*  
28: gb\_ges1:\*  
29: gb\_ges2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1310	90.4	2880	11	BC026195 Homo sapi
2	854	58.9	2584	11	AK030042 Mus muscu
3	852.4	58.8	1784	11	AK043543 Mus muscu
4	841.4	58.1	1201	13	BX365473 BX365473

5	809.8	55.9	1221	11	AK054237
6	714.4	49.3	937	13	BQ709169
7	692.4	47.8	932	13	EX328072
8	618.8	42.7	2239	11	AK030186
9	614	42.4	674	11	BC009414
10	570.2	39.4	671	10	BE388672
11	564.6	39.0	659	12	BE752438
12	510.6	35.2	767	12	BI544364
13	509.6	35.2	867	13	BU930538
14	504	34.8	845	14	CD106588
15	465.2	32.1	635	10	BE552498
16	462.2	31.9	849	13	EX428086
17	388	26.8	878	14	CD557726
18	356.2	24.6	1148	12	EM563364
19	354.4	24.5	609	10	BB666177
20	337.2	23.3	584	14	CD689119
21	336.4	23.2	1126	12	EM563363
22	325.2	22.4	713	12	EG744158
23	315.8	21.8	480	9	AI384010
24	313.2	21.6	628	10	BB616480
25	296.8	20.5	374	9	AI288640
26	273.2	18.9	776	12	BG718517
27	272.4	18.8	620	10	BB617055
28	240	16.6	616	10	BE262600
29	215.4	14.9	675	14	CD251322
30	210.8	14.5	612	13	BU674219
31	199	13.7	525	9	AA010796
32	178.8	12.3	471	12	BI032921
33	176.8	12.2	183	10	BF934127
34	171.6	11.8	609	13	BM197769
35	161.8	11.2	748	29	PT010C03R
36	154.4	10.7	500	12	BF190100
37	154.2	10.6	592	13	BW240639
38	149.8	10.3	586	29	CE484348
39	149.6	10.3	1522	11	BC015515
40	149.6	10.3	2206	29	AY403632
41	147.8	10.2	702	12	BI754945
42	147.8	10.2	913	13	BQ432111
43	147.8	10.2	1201	13	EX364188
44	146.4	10.1	2136	29	AY403634
45	146.4	10.1	3193	11	AK078809

#### ALIGNMENTS

RESULT 1  
BC026195  
LOCUS BC026195 2880 bp mRNA linear HTC 19-NOV-2003  
DEFINITION Homo sapiens NIMA (never in mitosis gene a)-related kinase 11,  
mRNA (CDNA clone IMAGE:4828434), containing frame-shift errors.

ACCESSION BC026195  
VERSION BC026195.1 GI:20072185  
KEYWORDS HTC  
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 2880)

REFERENCE  
AUTHORS  
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,  
Carninci, P., Prange, C., Raja, S.S., Loquellano, N.A., Peters, G.J.,  
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,  
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Y.,  
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,



Db	1566	TTTAGAGAACTCTGAGTCTCAGCTATTCCTCCATGAGCTCCACGAACTTGAATC	1625
Qy	1362	AAATTGAGAGATGCCACATCTCAGCTTGGATACCATG	1399
Db	1626	AAATTGAGAGATGCCACATCTCAGCTTGGATACCATG	1663
RESULT 2			
AK030042			
LOCUS		2584 bp	linear
DEFINITION		Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4932416N14 product:similar to MRNA, SIMILAR TO HUMAN HYPOPHYSICAL PROTEIN FLJ23495, COMPLETE CDS [Macaca fascicularis], full insert sequence.	
ACCESSION	AK030042	GI:26326024	
VERSION	AK030042.1		
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE			
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraktion of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	20499374		
PUBMED	11042159		
REFERENCE			
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuami, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, I., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		
MEDLINE	20530913		
PUBMED	11076861		
REFERENCE			
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.		
TITLE	Functional annotation of a full-length mouse cDNA collection		
JOURNAL	Nature 409, 685-690 (2001)		
REFERENCE			
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.		
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs		
JOURNAL	Nature 420, 563-573 (2002)		
REFERENCE			
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takada, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Direct Submission		
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			

Db	431	GCTCAGCTCTCTCCAGGCTCCATCATCTCTGCCATCTGTGAGATTCATGCTCAAGCTTCATG	490
Qy	298	GAGCAAGATAAATTTCTGCTATTTACGAGGAGTACTGTGAGGGCCGAGATCTGACGATATAA	357
Db	491	GACGAGGAGACTTTTGGCTATTTACGGAATACGTTGAGGGCCGAGATCTGACTATAGA	550
Qy	358	ATTCAGGAATATAAACAAGCTGGAATAATCTTCCAGAAAATCAATATATAGAAATGGTTT	417
Db	551	ATCCAGGAATATAAAGAAGCTGGAATAATCTTCCAGAAAATCAATATATAGAAATGGTTT	610
Qy	418	ATCCAGCTGCTGCTGGAGTTGACTACATGCATGAGAGGAGATCTTTCATCCAGACTTA	477
Db	611	ATCCAGCTGCTGCTGGAGTTGACTACATGCATGAGAGGAGATCTTTCATCCAGACTTG	670
Qy	478	AATCAAGAATATATTTCTGAAAATTAATCTCCTTAAATTTGGAGATTTTGGAGTTTCT	537
Db	671	AAATCAAGAATATATTTCTGAAAATTAATCTCCTTAAATTTGGAGATTTTGGAGTTTCT	730
Qy	538	CGACTTCTAATGGATCTCTGCTGACCTGGCCACAACTTTTAACTGGAATCTCCCATTTATG	597
Db	731	CGCTGCTAATGGTTTCACTGAGCTGGCTGACACTCTAACCGGACCCCATTTATG	790
Qy	598	AGTCTGAGGCTCTGAAACCAAGGCTATGACACAAAATGCGACATCTGGTCACTGGCA	657
Db	791	AGTCCGAGGCTCTGAAACCAAGGCTATGACACAAAATGCGACATCTGGTCACTGGCA	850
Qy	658	TGCATTTTGTAGATGCTGCTCATGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	717
Db	851	TGCATTTTGTAGATGCTGCTCATGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	910
Qy	718	ATTGTTTAAATAATTTGTAAGGTGACACACCTTCTCTCCCTGAGAGATATCAAAAGAA	777
Db	911	GTGGTTTGAATATTTGTAAGGTGAAACACCTTCTCTCCCTGAGAGATATCAACAGAA	970
Qy	778	CTAATGCCATCATGGAAGCATGTTGAAACAGAAATCTTCTTCAATAGACATCTGCTATC	837
Db	971	CTAACACCATCATGGAAGCATGTTGAAACAGAAATCTTCTTCAATAGACATCTGCTATC	1030
Qy	838	GAATTTTAAATAATTTGTAAGGTGACACACCTTCTCTCCCTGAGAGATATCAAAAGAA	897
Db	1031	GACATTTTAAAGCCCTTACATGAGAGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1090
Qy	898	GAATGACTCTGGAAGCAAAAATTTGGATGTCAGAGAGGCTGCTCATATTAATTAAT	957
Db	1091	GAGATGACACTGGAAGCAAAAATTTGGATGTCAGAGAGGCTGCTCATATTAATTAAT	1150
Qy	958	GCCATGCAAAAAGGATCCACCTGACGACTCTGAGGGCTGTCAGAGTACAGAAAATG	1017
Db	1151	GCGGTGAGAAAAGCTTCACTGCGACTCTGCAAGCTTCTGCAAGCTTCTGCTGCAAGAAAG	1210
Qy	1018	ACGCCAAGAAAAGGATCGCGCTGAGGAAGCTCCAGCGCTGATGAGAAAGCCAGGAAG	1077
Db	1211	ACTCCAGAGAAACGATGCTGCTGAGGAAGCTGAGGAGGCTGAGGAGGCTGAGGAGG	1270
Qy	1078	CTGAAAAGATTTGGGAGAAAATTAAGAAAATTAAGAAAATTAAGAAAATTAAGAAAATTA	1137
Db	1271	CTGAAAAGATTTGGGAGAAAATTAAGAAAATTAAGAAAATTAAGAAAATTAAGAAAATTA	1330
Qy	1138	TCTCGGAATTTTCCAGAGCTGAGTGTGATGTGATGTGATGTGATGTGATGTGATGTGATG	1179
Db	1331	TCCCGGAATGTGGGTCCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1372
RESULT 3			
AK043543			
LOCUS	1784 bp mRNA linear HTC 19-SEP-2003		
DEFINITION	Mus musculus 10 days neonate cortex cDNA, RIKEN full-length enriched library, clone:A830007B13 product:similar to MRNA, SIMILAR TO HUMAN HYPOTHETICAL PROTEIN FLJ23495, COMPLETE CDS [Macaca fascicularis], full insert sequence.		
ACCESSION	AK043543		
VERSION	AK043543.1 GI:26335750		
KEYWORDS	HTC; CAP trapper.		

SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
AUTHORS	Carninci, P. and Hayashizaki, Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
MEDLINE	99279253
PUBMED	10349636
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PUBMED	11042159
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Hatake, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwake, S., Inoue, K., Togawa, Y., Izawa, M., Chata, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, O., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	2050913
PUBMED	11076861
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
REFERENCE	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	6 (bases 1 to 1784)
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, I., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sobue, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (15-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN) Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in RIKEN Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
FEATURES	Please visit our web site for further details.
source	URL: http://genome.gsc.riken.go.jp/
	URL: http://fantom.gsc.riken.go.jp/
	Location/Qualifiers
	1..1784

```
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:A830007B13"
/db_xref="MGI:2408755"
/db_xref="taxon:10090"
/clone="A830007B13"
/tissue_type="cortex"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 days neonate"
471..>1784
/notes="unnamed protein product; putative
similar to MENA, SIMILAR TO HUMAN HYPOTHETICAL PROTEIN
FLJ23495, COMPLETE CDS [Macaca fascicularis]
(SPTR|BA883539, evidence: FASTY, 72.4%ID, 100%length,
match=1806)"
/codon_start=1
/protein_id="BAC31576.1"
/db_xref="GI:26335751"
/translation="MLKFOSTAKCVGRRTVTPMPTALIRRYVYVQKLGSGFQV
YLVSKKAKPGBELVLEISVGEINPNETVQANVEAQLLRHLHPAIVRFHAFMEQ
ETFCITTEYCEGRDLDIRIOEYKEAGKFAENQIVWFQILGLVDYVHERRILHRDL
KSKNIFLKNLKIIGDFGSRLLMGSCLEATLTGAPHYMSPEALKHOGYDAKSDIWS
LACILVEMCLDHPAGSFLSVLNIVEGTPSLPDYPRELNTIMERMLNKPSSLR
PSAADILKAPYMEEOQLQLMCKYPMETLEDKNSVCCKEAAHTINAVQKLHLQAL
STQKTPRERWLRLQADERRARLKKIAENYKENDKRMQALRSRNSGVHAHVL
HELDERTUESLPOLPCLDLELPEPSLEDTIVDLGHEKRRKER"
ORIGIN
Query Match 58.8%; Score 852.4; DB 11; Length 1784;
Best Local Similarity 83.2%; Pred. No. 1.1e-192;
Matches 993; Conservative 0; Mismatches 196; Indels 3; Gaps 1;
QY 1 ATGCTGAATTCAGAGCAGCTAAGTGTGTG---AGTGGATCAACGCCATTTCCACT 57
DB 471 ATGCTGAATTCAGAGACTCGAAGTGTGTGGGAGAGAGCCACACAGTCATCCCATG 530
QY 58 TATCCAAAGACCTTGATTCAGAGATACGTGCTTCAACAAATCTTGGCAGTGGAGT 117
DB 531 TACCCGACGCTTTGATTCAGAGATATGTTCTTACAGAAATCTTGGCAGTGGAGT 590
QY 118 TTGGAATCTCTATCTGTTTTCAGACAAAGAACCCAGAGAGAGAGAAATTAAGGTA 177
DB 591 TTGGAATCTCTATCTGTTTTCAGACAAAGAACCCAGAGAGAGAGTAAAGGTA 650
QY 178 CTTAGGAATATCTGTTGGAGACTAATCAATCAATGAACTGTACAGCCCAATTTGAA 237
DB 651 CTGAGGAATCTCTGTTGGCAATTAATCCAAATGAGACCGCTGAGGCCCAATGTGAA 710
QY 238 GCCCACTCTCTCCAACTGACACCCAGCCATTTCAAGTTCATGCAAGTTTGTG 297
DB 711 GTCAGCTCTCTCCAGCTCCATCATCTGCCATGTCAGATTCCATGCAAGTTCTATG 770
QY 298 GAGCAAGATAATTCGCAATATACGGAGTACTGTGAGGCGGAGATCTGGACGATATAA 357
DB 771 GAGCAGAGACTTTTGTGATATACGGAAATCTGTGAGGCGGAGATCTGAGCTATAGA 830
QY 358 ATTCAAGATATAAACAAGCTGGAATAATCTTTCAGAAATCAATAATGAAGTTT 417
DB 831 ATCCAGGATATAAGAACTGGAAAGTCTTCGCTGAAATCAGATAGTAAGTTT 890
QY 418 ATCCAGCTCTGCTGGAGTTGACTATCATGATGAGAGAGAGGATATCTTCATCGAGCTTA 477
DB 891 ATCCAGCTGTTGCTGGAGTTGACTATCATGATGAGAGAGAGATATCTTCATCGAGCTG 950
QY 478 AGTCAAGAAATGATTTCTGAAAAATATCTCCITTAATTTGGAGATTTTGGAGTTCT 537
DB 951 AAATCAAGAAATATATTTCTGAAAAATATCTACTCAAAATCGGGATTTTGGAGTTCT 1010
QY 538 CGACTTCTAATGGGATCTGTGACCTGGCCACAACTTTTAACCTGGAACTCCCATTTATG 597
DB 1011 CGGCTCTAATGGGTTTCATGTGAGCTGGCTACAACTTAACCGGGGCCCCCATTTATG 1070
```



TITLE  
JOURNAL

Direct Submission  
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)

COMMENT  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/  
Location/Qualifiers  
1. .1221  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:E230038G21"  
/db\_xref="MGI:2426244"  
/db\_xref="taxon:10090"  
/clone="E230038G21"  
/sex="female"  
/tissue type="oviduct"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="2 days pregnant adult"  
126. .>1220  
/note="unnamed protein product; putative  
similar to MRNA, SIMILAR TO HUMAN HYPOTHETICAL PROTEIN  
FLJ23495, COMPLETE CDS [Macaca fascicularis]  
(SPTRI|BAB83539, evidence: FASTY, 72.4%ID, 100%length,  
match=1806)"  
/codon\_start=1  
/protein\_id="BAC35699.1"  
/db\_xref="GI:26344085"  
/translation="MLKQETAKVGRVETIPYPTALIRYVLOQLKSGSGFTV  
YLVSKRKPGEDELIVKISVGEINPNETVOANVEAQLSLHHPAIVRPHASFMQE  
EYFCITTECEGRDIDYIOEYKEAGVFAENOIVFWFTOLLGLVDYVHERILHRDL  
KSNIFLKNLAKIGDPGVSRLIMSGSELATLTGTPTPHYSPEALKHOGYDAKSDIWS  
LACILYEMCLOHAEAGSSFLSVLNVIEGKTPSLPDYPRELNTIMERMLNKSRLR  
PSAAILKAPYMEEQULWCKPYEMTLEDKNSVCQKEAAHTINAVQKRLHLQTLQAL  
SDTKQTPTRMWRMLRKLQADDERARLKKIA"

CDS

Query Match 55.9%; Score 809.8; DB 11; Length 1221;  
Best Local Similarity 84.4%; Pred. No. 1 6e-182;  
Matches 923; Conservative 0; Mismatches 167; Indels 3; Gaps 1;

QY 1 ATGCTGAAATCCAGAGGAGCTAAGTGTGTG---AGTGGATCAACAGCCATTCCACT 57  
126 ATGCTGAAATCCAGAGAGCTGGAAGTGTGTGGCAGAAGACCCACAGTCATCCCATG 185

QY 58 TATCCAAAGACCTTGATTCGAAGAATACGTCCTTCAACAAAATTCGCAAGTGAAGT 117  
186 TACCCAGCCGCTTTGATTCGAAGAATATGTCCTTCAGAGAACTTGGCAGTGAAGT 245

QY 118 TTGGAACTGTCTATCTGTTTCAGACAGAAGCCAAACGAGGAGGAGAAATTAAGGTA 177  
246 TTGGAACTGTCTATCTGTTTCAGACAGAAGCCAAACCTGGAGAGGAGCTAAGGTA 305

QY 178 CTTAAGGAAATATCTGTTGAGAACTTAATCCAATGAAACTGTACAGCCCAATTGGAA 237  
306 CTGAAGGAAATCTCTGTTGGCGAAATTAATCCAATGAGACCGTGGAGCCCAATGGAA 365

QY 238 GCCCAACTCTCTCCAAAGCTGGACCCAGCCATGTCGAAGTTCATCAAGTTTGTG 297  
366 GCTCAGCTCTCTCCAGGCTCCATCATCTGCGCATTTGTGATTCATCAAGCTTCATG 425

QY 298 GACACAGATAATTTCTGCATTATCAACGAGTACTGTGAGGCGGAGATCTGACGATAA 357

Db QY 426 GAGCAGGAGACTTTTTCATTATCACGGAATACCTGTGAGGSCCGAGATCTCGACTATAGA 485

QY 358 ATTGAGGAATATAAAGCTGGAAAATCTTCCAGAAAATCAAAATTAATAGATGTTT 417

Db 486 ATCCAGGAATATAAAGAGCTGGGAAAAGTCTTCGCTGAAAATCAGATAGTAGATGTTT 545

QY 418 ATCCAGCTGCTGCTGGGAGTTGACTACATGATCAGAGGAGGAGTACTTTCATCGAGACTTA 477

Db 546 ATCCAGCTGTTGCTTGGAGTTGACTACATGATCAGAGGAGGAGTACTTTCATCGAGACTTG 605

QY 478 AAGTCAAGAAATGATTTTCTGAAAATAATTCCTTAAAATTGAGAGATTTTGGAGTTTCT 537

Db 606 AAATCAAGAAATATATTTCTGAAAATAATCTACTCAAAATCGGGGATTTTGGAGTTTCT 665

QY 538 CGACTTCTAATGGGATCTGTGACTGGCCACAACTTTAACTGGAACCTCCCAATTATATG 597

Db 666 CGGCTGCTAATGGGTTCAITGAGCTGGCTACAACTCTAACCGGAGCCCCCAATTATATG 725

QY 598 AGTCCTGAGGCTCTGAAACACCAAGGCTATGACACAAAGTCGACATCTGGTCACTGGCA 657

Db 756 AGTCGAGGAGGCTTGAAGCAAGGCTATGATGCCAAGTCTGACATCTGGTCACTGGCA 785

QY 658 TGCATTTTGTATGATGATGCTGTCATGATGATGATGATGATGATGATGATGATGATGAT 717

Db 786 TGCATTTTATATGATGATGTTGTGCTGGATCATGATGATGATGATGATGATGATGATGAT 845

QY 718 ATTGTTTTAAAAATTTGTAAGGTCAACACACTTCTCTCCCTGAGAGATATCCAAAAGAA 777

Db 846 GTGTTTTGATATTTGTAAGGTAAACACCTTCACTCCCGGACAGATATCCACGAGAA 905

QY 778 CTAATGTCATCATGGAAGCATGTTGAAAGAAATCTTTCATTAAGACCATCTGCTATC 837

Db 906 CTAACACCATCATGGAAGCATGTTGAAAGAAAGTCTTCAATTTGAGACCGCTCGGCTGCA 965

QY 838 GAAATTTTAAAAATCCCTTACCTTGTGAGCAGCTACAGAACCTTAATGTGTAGATATTC 897

Db 966 GACATTTTAAAAAGCCCTTACATGGAAGAGCAGCTTCACTCTCTGTGTGTAAATACCCA 1025

QY 898 GAAATGACTCTGGAAGCAAAAATTTGGATTTGTCAGAGGAGGCTGCTCATATTAATTAAT 957

Db 1026 GAGATGACACTGGAAGACAGAACTCAGTTTGTGTCAGAGGAGGCTGCTCATACCATTAAT 1085

QY 958 GCCATGCAAAAAGGATCCACTCGACAGCTCTGAGGCACTGTGAGAGTACAGAAATG 1017

Db 1086 GCCGTGCAAAAAGGATTTCACTCGACACTCTGCAAGCTTGTCTGACACCCAGAAACG 1145

QY 1018 ACGCAAGAGAAAGGATGCGGCTGAGGAAGCTCCAGGCGGCTGATGAGAAACCCAGGAAG 1077

Db 1146 ACTCCAGAGAAACGAGTGTGCTGAGGAAGCTGCAAGCAGCCGATGAGAGAGCCAGGAG 1205

QY 1078 CTGAAAAGAGATTG 1090

Db 1206 CTGAAAAGAGATTG 1218

RESULT 6  
BQ709169  
LOCUS  
DEFINITION BQ709169 337 bp mRNA linear EST 16-JUL-2002  
5' mRNA sequence.  
ACCESSION BQ709169  
VERSION BQ709169.1 GI:21848068  
KEYWORDS EST  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 937)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov

Tissue Procurement: Dr. Mark Watson  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov

Plate: L1CM2518 row: k column: 12

High quality sequence start: 4

High quality sequence stop: 638.

## FEATURES

source

1..937  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6301955"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_113"  
/note="Organ: spleen; Vector: pOT37; Site\_1: XhoI; Site\_2:  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGCAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC Library."

## ORIGIN

Query Match 49.3%; Score 714.4; DB 13; Length 937;  
Best Local Similarity 96.9%; Pred. No. 9.2e-160;  
Matches 760; Conservative 0; Mismatches 21; Indels 3; Gaps 3;  
QY 1 ATGCTGAATTCGAAGGAGGAGTAACTGTCGAGTGGATCAACAGCCATTTCCACTTAT 60  
DB 153 ATGCTGAATTCGAAGGAGGAGTAACTGTCGAGTGGATCAACAGCCATTTCCACTTAT 212  
QY 61 CCAAGACCTTGATTCGAAGAAGATAGTCTCTTCAACAAAACTTGGCAGTGAAGTTT 120  
DB 213 CCAAGACCTTGATTCGAAGAAGATAGTCTCTTCAACAAAACTTGGCAGTGAAGTTT 272  
QY 121 GGAAGTCTATCTGTTTCAGACGAAGAGCCAAACGAGGAGAGGAATTAAGTACTT 180  
DB 273 GGAAGTCTATCTGTTTCAGACGAAGAGCCAAACGAGGAGGAATTAAGTACTT 332  
QY 181 AAGGAATATCTGTGGAGAACTTAATCCAAATGAACTGTACAGGCCAATTTGGAAGCC 240  
DB 333 AAGGAATATCTGTGGAGAACTTAATCCAAATGAACTGTACAGGCCAATTTGGAAGCC 392  
QY 241 CAATCTCTCCAGGTGGACCCAGCCAGCCATTTCAAGTTCATCAAGTTTGTGGAG 300  
DB 393 CAATCTCTCCAGGTGGACCCAGCCAGCCATTTCAAGTTCATCAAGTTTGTGGAG 452  
QY 301 CAAGATAATTTCTGCATTATCAGGAGTACTGTGAGGCCGAGATCTGCAGCATATAAAT 360  
DB 453 CAAGATAATTTCTGCATTATCAGGAGTACTGTGAGGCCGAGATCTGCAGCATATAAAT 512  
QY 361 CAGGAATATAACAAGCTGGAATAATCTTTCCAGAAAAATCAAAATATAGAATGGTTTATC 420  
DB 513 CAGGAATATAACAAGCTGGAATAATCTTTCCAGAAAAATCAAAATATAGAATGGTTTATC 572  
QY 421 CAGCTGCTGGAGTTGATCATGATGAGAGGAGATCTTCATCGAGACTTAAG 480  
DB 573 CAGCTGCTGGAGTTGATCATGATGAGAGGAGATCTTCATCGAGACTTAAG 632  
QY 481 TCAAGAATCTATTCTGAAAAATATCTCTTAAAAATGGAGATTTTGGAGTTTCTCGA 540  
DB 633 TCAAGAATCTATTCTGAAAAATATCTCTTAAAAATGGAGATTTTGGAGTTTCTCGA 692  
QY 541 CTTCTAATGGAT-CTGTGACCTGGCCACAACCTTTAACTGGAACTCCCACTTATATGAG 599  
DB 693 CTTCTAATGGATCCCTGTGACCTGGCCACAACCTTTAACTGGAACTCCCACTTATATGAG 752  
QY 600 TCCTGAGGCTCTGAACACCAAGGCTATGACACAAAGTCGACATCT-GGTCACTGGCAT 658

DB 753 TCCTGAGGCTCTGAACCCCAAGGCTATGACACAAAGTCGACATCTGGGTCACTGCAT 812  
QY 659 GCATTTTGTATGAGATGCTGTCATGATCATGATCATGATCGGTGGCTCCAAATTTCTATCCA 718  
DB 813 GCATTTTGTATGAGATGCTGTCATGATCATGATCATGATCGGTGGCTCCAAATTTCTATCCA 872  
QY 719 TTGTTTAAAAAATTTGTTGAAGTGAACA-CACCTTCTCTCCTCTGAGATATCCAAAAGAA 777  
DB 873 ATGTTTAAAAAATGTTGAAGGGGACACCACTTCCCTCCTCGGAAATATTCAAAAGAA 932  
QY 778 CTAA 781  
DB 933 CTAA 936

## RESULT 7

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

932 bp mRNA linear EST 01-MAY-2003  
BX328072 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED  
Homo sapiens cDNA clone CS00J004YB13 5-PRIME, mRNA sequence.  
BX328072  
BX328072.1 GI:30309740  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Li.W.B., Gruber,C., Jesse,J. and Polayes,D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 2479. For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0BAG041ZF11\_CS03926\_1acluster=2479.f.  
Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0BAG041ZF11\_CS03926\_1.

## FEATURES

source

1..932  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS00J004YB13"  
/cell\_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"  
/cell\_line="JURKAT"  
/clone\_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT  
10-NORMALIZED"  
/note="First strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

Query Match 47.8%; Score 692.4; DB 13; Length 932;  
Best Local Similarity 98.7%; Pred. No. 1.7e-154;  
Matches 708; Conservative 0; Mismatches 7; Indels 2; Gaps 1;  
QY 683 TGAATCATGCAATCGCTCGCTCCAAATTTCTTATCCATTTGTTTAAAAAATTTGTTGAAGTG 742  
DB 11 TGNATCATGCAATCGCTCGCTCCCA--TTTCTTATCTCTTGTTTAAAAAATTTGTTGAAGTG 68  
QY 743 ACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTTAATGCCATCATCGAAGCATGT 802  
DB 69 ACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTTAATGCCATCATCGAAGCATGT 128  
QY 803 TGAACAAAGATCTCTTATTAAGACCATCTGCTATCGAAATTTTAAAAAATCCCTTACCTTG 862



```

Db      129  TGAACAGAACTCTTCATTAAAGACCATCTGCTATCGAAATTTTAAATAACCTTACCTTG 188
QY      863  ATGAGCAGCTACAGACCTTAATGTGTAGATATTAGAAATGACTCTGGAGACAAAAATT 922
Db      189  ATGAGCAGCTACAGACCTTAATGTGTAGATATTAGAAATGACTCTGGAGACAAAAATT 248
QY      923  TGGATTGTTCAGAGGAGGCTGCTCATATTAATTAATCCATGCAAAAAGATCCACCTGC 982
Db      249  TGGATTGTTCAGAGGAGGCTGCTCATATTAATTAATCCATGCAAAAAGATCCACCTGC 308
QY      983  AGACTCTGAGGGCACTGTTCAGAAAGTACAGAAATGACGCCAAGAGAAAGATCGCGCTGA 1042
Db      309  AGACTCTGAGGGCACTGTTCAGAAAGTACAGAAATGACGCCAAGAGAAAGATCGCGCTGA 368
QY      1043  GGAAGCTCCAGGGCTGTATGAGAAAGCCAGGAGCTGAAAAGATTTGGAGAAAT 1102
Db      369  GGAAGCTCCAGGGCTGTATGAGAAAGCCAGGAGCTGAAAAGATTTGGAGAAAT 428
QY      1103  ATGAAGAAATAGCAAAACGAATCCAGAAATTTGAGATCTCGAACTTTTCAGCAGCTGAGTG 1162
Db      429  ATGAAGAAATAGCAAAACGAATCCAGAAATTTGAGATCTCGAACTTTTCAGCAGCTGAGTG 488
QY      1163  TTGATCTACTCCATGAAAACACATTTAAAGGAATGGAAGAAAGGAGGAGCACTG 1222
Db      489  TTGATCTACTCCATGAAAACACATTTAAAGGAATGGAAGAAAGGAGGAGCACTG 548
QY      1223  AGGGAAGACTTCTTCTGTTCAACCCAGGAGGATCAAGAGAGTGGCAGGAGGAGGAG 1282
Db      549  AGGGAAGACTTCTTCTGTTCAACCCAGGAGGATCAAGAGAGTGGCAGGAGGAGGAG 608
QY      1283  AGGAATCTGATGAAACCACTTTAGAGAACTGCTGAGTCTGAGCTATTCCTTCATGG 1342
Db      609  AGGAATCTGATGAAACCACTTTAGAGAACTGCTGAGTCTGAGCTATTCCTTCATGG 668
QY      1343  ACCTCCAGCAACTTGAATCAATTTGAGAGATGCCACATCTGACCTTGGATACCATG 1399
Db      669  ACCTCCAGCAACTTGAATCAATTTGAGAGATGCCACATCTGACCTTGGATACCATG 725

RESULT 8
AK030186 2239 bp mRNA linear HTC 18-SP-2003
LOCUS Mus musculus adult male testis cDNA, RIKEN full-length enriched
DEFINITION library, clone:4933413H15 product:similar to MRNA, SIMILAR TO HUMAN
HYPOTHETICAL PROTEIN FL023495, COMPLETE CDS [Macaca fascicularis],
full insert sequence.
ACCESSION AK030186
VERSION 1 GI:26081701
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Takahashi, Y., Itoh, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
JOURNAL Prepare full-length cDNA libraries for rapid discovery of new genes
MEDLINE Genome Res. 10 (10), 1617-1630 (2000)
PUBMED 20499374
REFERENCE 2
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akizawa, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, I., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,

```

```

Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
MEDLINE 11076861
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, P., Takaku-Akaira, S.,
Takeda, Y., Tanaka, I., Tomaru, A., Toyota, I., Yasunishi, A.,
Muramatsu, M., and Hayashizaki, Y.
Direct Submission
SUBMITTED (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Sushiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.
FEATURES
source
1..2239
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="FANTOM DB:4933413H15"
/db_xref="MGI:2392882"
/db_xref="taxon:10090"
/clone="4933413H15"
/sex="male"
/tissue_type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
misc_feature
235..1598
/note="putative
similar to MSNA, SIMILAR TO HUMAN HYPOTHETICAL PROTEIN
FL023495, COMPLETE CDS [Macaca fascicularis]
(SPTR|BAB83539, evidence: PASTY, 72.4%ID, 100%length,
match=1806)"
2221..2226
/note="putative"
2239
/note="putative"
polyA_signal
2239
polyA_site
2239
ORIGIN

```

Query Match		42.7%; Score 618.8; DB 11; Length 2239;
Best Local Similarity		74.0%; Pred. No. 7e-137;
Matches 875; Conservative		0; Mismatches 177; Indels 130; Gaps 2;
Qy	1	ATGCTGAAATTCGAAGGAGCTAAGTGTGTG---AGTGGATCAACAGCCATTTCCACT 57
Db	235	ATGCTGAAATTCGAAGAGACTGCGAAGTGTGTGGCAGAAGACCCACAGCTCATCCCATG 294
Qy	58	TATCCAAAGACCTTGATTCGAAGAGATAGTCTTCAACAAAACCTTCGAGTGGAGCT 117
Db	295	TACCCGACCTTTGATTTGCAAGAGATATGTTCTTCAGCAAACTTCGCGTGGAGT 354
Qy	118	TTTGGAACTCTATCTGTGTTTCAGCAAGAAAGCCAAAGCAGGAGAGGAATTAAGGTA 177
Db	355	TTTGGAACTCTATCTGTGTTTCAGCAAGAAAGCCAAAGCAGGAGAGGAATTAAGGTA 414
Qy	178	CTTAAGGAATATCTGTGAGAACTTAATCCAAATGAACCTGTACAGGCCAATTTGGA 237
Db	415	CTGAAGGAATCTCTGTGTGCGAATTAATCCAAATGAGACCGTGCAGGCCAATGTGGA 474
Qy	238	GCCCAACTCTCTCCAAAGCTGGACCCAGCCATTTGCAAGTTCCATGCAAGTTTGTG 297
Db	475	GCTCAGCTCTCTCCAGGCTCCATCTCCGCTATGTCAGATTCATGCAAGCTTCATG 534
Qy	298	GAGCAAGATATTTCTGCATATACGAGTACTGTGAGGCGCAGATCTGACGATAAA 357
Db	535	GAGCAGGAGACTTTTTCATTATCACGGAATACTGTGAGGCGCAGATCTGGAATAGA 594
Qy	358	ATTGAGGAATATAACCAAGCTGGAATAATCTTCCAGAAATCAATATAGAAATGGTTT 417
Db	595	ATCCAGGAATATAAGAACTGGAAAGTCTTCTGTAATTCAGATAGTAATGGTTT 554
Qy	418	ATCCAGCTGCTGTGGAGTGTACTATCATCATGAGAGGAGGATCTTATCATGAGACTTA 477
Db	655	ATCCAGCTGTTGCTGTGAGTGTACTATCATCATGAGAGGAGGATCTTATCATGAGACTTG 714
Qy	478	AAGTCAAGAGATGATTTCTGAAATAATCTCCTTAAATTTGGAGATTTGAGTTTCT 537
Db	715	AAATCAAGAGATATATTTCTGAAATAATCTTACTCAAAAT----- 755
Qy	538	CGACTTCTAATGGATCCTGTGACTGCGCCACAACTTTAACTGGAACTCCCATATATG 597
Db	756	----- 755
Qy	598	AGTCTGAGGCTGTGAACACCAAGGCTATGACACAAAGTCGACATCTGGTCACTGGCA 657
Db	756	-----CGGTCACTGGCA 767
Qy	658	TGCATTTTGTATGATGTGCTGCATGAATCATGCAATTCGCTGGCTCCCAATTTCTTATCC 717
Db	768	TGCATTTTATATGATGTGCTGCATGATGCAATTTGCTGCTCCAGTTTCTTGTCT 827
Qy	718	ATTGTTTAAATAATGTTGAAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAA 777
Db	828	GTGTTTGAATATTTGTTGAAGTAAACACCTTCACTCCCGACAGATATCCACAGAA 887
Qy	778	CTAATGCCATCATGGAAGCATGTTCAACAGAACTCTTCAATTAAGCATCTGTATC 837
Db	888	CTAAACACCATCATGGAACGATGTTGAACAGAGTCTTCAATTAAGCATCTGTATC 947
Qy	838	GAAATTTTAAATAATCCCTTACCTTGTATGAGCAGCTACAGAACTTAATGTGATATTC 897
Db	948	GACATTTTAAAGGCCCTTACATGGAAGAGCAGCTTCAAGTCTCTGTATGTAATACCCA 1007
Qy	898	GAAATGACTCTGGAAGCAAAAATTTGGATGTGAGAGGAGCTGCTCATATAATTAAT 957
Db	1008	GAGATGACTCTGGAAGCAAAAATTTGGATGTGAGAGGAGCTGCTCATATAATTAAT 1067
Qy	958	GCCATGCAAAAAGGATCCACTGACACTCTGAGGCACTGTGAGAGTACAGAAATG 1017
Db	1068	GCCGTGCAAAAAGGATCCACTGACACTCTGAGGCACTGTGAGAGTACAGAAATG 1127
Qy	1018	ACGCCAAGAAAGGATCGGCTGAGGAGCTCCAGCGGCTGATGAGAAAGCCAGGAG 1077

Db	1128	ACTCCAGAGAACGATGTGTGCTGAGGAAGCTGCGAGGAGCCGATGAGAGCCAGAGG 1187
Qy	1078	CTGAAAAGATTTGGAGAAAATATGAAGAAATAGCAAAACGAATGCAAGATTTGAGA 1137
Db	1188	CTGAAAAGATTTGCTGAAGAAACTATGAAGAAATGACAAAGGATGCGAGCCCTGAGA 1247
Qy	1138	TCTCGGAACITTCAGCAGCTGAGTGTGTGATGTACTCCATGAA 1179
Db	1248	TCCCGAATGTGGGTCCGTGCATGCCCATGTGCTCCATGAA 1289
RESULT 9		
BC009414		
LOCUS		
DEFINITION		
Homo sapiens NIMA (never in mitosis gene a)-related kinase 11, mRNA (cDNA clone IMAGE:3613037), with apparent retained intron.		
ACCESSION		
BC009414		
VERSION		
BC009414.1 GI:14714382		
KEYWORDS		
HTC.		
SOURCE		
Homo sapiens (human)		
ORGANISM		
Homo sapiens		
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
1 (bases 1 to 674)		
REFERENCE		
AUTHORS		
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altshul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheer, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwen, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Pahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalhus, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.		
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		
2388257		
MEDLINE		
12477932		
2 (bases 1 to 674)		
JOURNAL		
Strausberg, R.		
Direct Submission		
Submitted (19-JUN-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 1A03, Bethesda, MD 20892-2590, USA		
NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>		
Contact: MGC help desk		
Email: <a href="mailto:cgapbs-re@mail.nih.gov">cgapbs-re@mail.nih.gov</a>		
Tissue Procurement: ATCC		
CDNA Library Preparation: Rubin Laboratory		
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
BC Sequencing by: Genome Sequence Centre,		
BC Cancer Agency, Vancouver, BC, Canada		
info@bcgsc.bc.ca		
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.		
REMARK		
COMMENT		

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAL Plate: 23 Row: p Column: 9  
 This clone has the following problem: retained intron.

## FEATURES

source  
 1..674  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3613037"  
 /tissue\_type="uterus, endometrium adenocarcinoma"  
 /clone\_lib="NIH MGC 44"  
 /lab\_host="DH10B-R"  
 /note="Vector: pOTB7"

## ORIGIN

Query Match 42.4%; Score 614; DB 11; Length 674;  
 Best Local Similarity 98.1%; Pred. No. 8.4e-136;  
 Matches 632; Conservative 0; Mismatches 10; Indels 2; Gaps 1;

QY 170 TAAAGGTACTTAAGGAATATCTGTGGAGACTAAATCCCAATGAACTGTACAGGCCA 229  
 DB 1 TAAAGGTACTTAAGGAATATCTGTGGAGACTAAATCCCAATGAACTGTACAGGCCA 60  
 QY 230 ATTTCGAGCCCACTCTCTCCAAAGCTGGACCCAGCCATTTGTCAGTTCCATGCAA 289  
 DB 61 ATTTCGAGCCCACTCTCTCCAAAGCTGGACCCAGCCATTTGTCAGTTCCATGCAA 120  
 QY 290 GTTTTGTGAGCAAGATATTTCTGCATATCACGGAGTACTGTGAGGCCGAGATCTGG 349  
 DB 121 GTTTTGTGAGCAAGATATTTCTGCATATCACGGAGTACTGTGAGGCCGAGATCTGG 180  
 QY 350 ACGATAAATTCAGGAATATAAACAAGCTGGAAAAATCTTCCAGAAAAATCAATAATAG 409  
 DB 181 ACGATAAATTCAGGAATATAAACAAGCTGGAAAAATCTTCCAGAAAAATCAATAATAG 240  
 QY 410 AATGTTTATCAGCTGTCTGGAGTTGACTACATCATGATGAGAGGAGTACTTCATC 469  
 DB 241 AATGTTTATCAGCTGTCTGGAGTTGACTACATCATGATGAGAGGAGTACTTCATC 300  
 QY 470 GAGACTTAAAGTCAAGATGTATTTCTGAAATAATCTCTTAAATTTGAGATTTTG 529  
 DB 301 GAGACTTAAAGTCAAGATGTATTTCTGAAATAATCTCTTAAATTTGAGATTTTG 360  
 QY 530 GAGTTTCTCGACTCTTAATGGGATCTGTGACCTGGCCACAACTTTAACTGGAACTCCC 589  
 DB 361 GAGTTTCTCGACTCTTAATGGGATCTGTGACCTGGCCACAACTTTAACTGGAACTCCC 420  
 QY 590 ATTATATGAGTCTGAGGCTCTGAACACCAAGGCTATGACACAAAGTCGGACATCTGGT 649  
 DB 421 ATTATATGAGTCTGAGGCTCTGAACACCAAGGCTATG--ACAAAGTCGGACATCTGGT 478  
 QY 650 CACTGGCATGCATTTTGTATGATGTGCTGCATGAATCATGATTCGCTGGCTCCAATT 709  
 DB 479 CACTGGCATGCATTTTGTATGATGTGCTGCATGAATCATGATTCGCTGGCTCCAATT 538  
 QY 710 TCTTATCCATTTGTTTAAAAATTTGTAAGGTGACACACCTTCTCTCCCTGAGATATC 769  
 DB 539 TCTTATCCATTTGTTTAAAAATTTGTAAGGTGACACACCTTCTCTCCCTGAGATATC 598  
 QY 770 CAAAGCACTAAATGCCATGCAAGCATGTTGAAACAAGAT 813  
 DB 599 CAAAGCACTAAATGCCATGCAAGCATGTTGAAACAAGAT 642

## RESULT 10

BE388672  
 LOCUS 601286319F1 NIH\_MGC\_44 Homo sapiens cDNA clone IMAGE:3613037 5',  
 DEFINITION mRNA sequence.  
 BE388672  
 ACCESSION BE388672.1 GI:9334037  
 VERSION  
 KEYWORDS EST.

## SOURCE

ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 671)

## AUTHORS

NIH-MGC <http://mgs.nci.nih.gov/>.

## TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

## JOURNAL

Unpublished (1999)

## COMMENT

Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LICM273 row: e column: 06  
 High quality sequence stop: 602.

## FEATURES

source

1..671  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3613037"  
 /tissue\_type="endometrium, adenocarcinoma cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH MGC 44"  
 /note="Organ: uterus; Vector: pOTB7; Site 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGCAAGAG(G). Library constructed by Ling Hong  
 in the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies)."

## ORIGIN

Query Match 39.4%; Score 570.2; DB 10; Length 671;  
 Best Local Similarity 97.1%; Pred. No. 2.4e-125;  
 Matches 612; Conservative 0; Mismatches 14; Indels 4; Gaps 3;

QY 189 ATCTGTTGGAGCACTAAATCCAAATGAACTGTACAGGCCAAATTGGAAGCCCAACTCTCT 248  
 DB 1 ATCTGTTGGAGCACTAAATCCAAATGAACTGTACAGGCCAAATTGGAAGCCCAACTCTCT 60  
 QY 249 CTCGAAGCTGGACCAACCCAGCCATTGTCAAGTTCCATGCAAGTTTGTGAGCAAGATAA 308  
 DB 61 CTCGAAGCTGGACCAACCCAGCCATTGTCAAGTTCCATGCAAG-TTCTGGAGCAAGATAA 119  
 QY 309 TTTCTGCATTTACCGAGTACTGTGAGGCCCGAGATCTGGACGATAAAATTCAGGAATA 368  
 DB 120 TTTCTGCATTTACCGAGTACTGTGAGGCCCGAGATCTGGACGATAAAATTCAGGAATA 179  
 QY 369 TAAACAAGCTGGAATAATCTTCCAGAAAAATCAATAATAGAAATGTTTATCCAGCTGCT 428  
 DB 180 TAAACAAGCTGGAATAATCTTCCAGAAAAATCAATAATAGAAATGTTTATCCAGCTGCT 239  
 QY 429 GCTGGAGTTGACTACATGATGATGAGAGGAGTACTTCATCGACACTTAAAGTCAAGAA 488  
 DB 240 GCTGGAGTTGACTACATGATGATGAGAGGAGTACTTCATCGACACTTAAAGTCAAGAA 299  
 QY 489 TGTATTTCTGAAAAATAATCTCTTAAAAATTTGGAGATTTTGGAGTTTCTCGACTCTAAT 548  
 DB 300 TGTATTTCTGAAAAATAATCTCTTAAAAATTTGGAGATTTTGGAGTTTCTCGACTCTAAT 359  
 QY 549 GGGATCTCTGTGACTGGCCCAACTTTTAACTGGAACTCCCATTAATATGAGTCTCTGAGGC 608  
 DB 360 GGGATCTCTGTGACTGGCCCAACTTTTAACTGGAACTCCCATTAATATGAGTCTCTGAGGC 419  
 QY 609 TCTGAACACCAAGCTATGACACAAAGTGGACATCTGTCACCTGCATGCACTTTTGTA 668  
 DB 420 TCTGAACACCAAGGCTATG--ACAAAGTCGGACATCTGGTCACTGCAATGCA-TTTGTA 476

QY 669 TGAGATGCTGCATGATCATGCAATTCGCTGCTCCAAATTCCTTATCCATTGTTTAAA 728  
 Db 477 TGAGATGCTGCATGATCATGCAATTCGCTGCTCCAAATTCCTTATCCATTGTTTAAA 536  
 QY 729 AATGTTTGAAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTAAATGCCAT 788  
 Db 537 AATCGTTGAAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTAAATGCCAT 596  
 QY 789 CATGGAAGCATGTTGAACAAGATCCTTC 818  
 Db 597 CATGGAAGCATGTTGAACAAGATGCTC 626

RESULT 11  
 BG752438 659 bp mRNA linear EST 15-MAY-2001  
 LOCUS 602730927F1 NIH\_MGC\_43 Homo sapiens cDNA clone IMAGE:4874479 5',  
 mRNA sequence.  
 ACCESSION BG752438  
 VERSION BG752438.1 GI:14063091  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE NIH-MGC http://mgs.nci.nih.gov/.  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished (1999)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: cgapbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLC1754 row: e column: 08  
 High quality sequence stop: 657.  
 Location/Qualifiers  
 1..659  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4874479"  
 /tissue\_type="normal pigmented retinal epithelium"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_43"  
 /note="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2:  
 EcoRI; cDNA made by oligo-dT priming. Directionally  
 cloned into EcoRI/XhoI sites using the following 5'  
 adaptor: GGACGAG(G). Library constructed by Ling Hong  
 in the laboratory of Gerald M. Rubin (University of  
 California, Berkeley) using ZAP-cDNA synthesis kit  
 (Stratagene) and Superscript II RT (Life Technologies).  
 Note: this is a NIH\_MGC Library. |"

ORIGIN  
 Query Match 39.0%; Score 564.6; DB 12; Length 659;  
 Best Local Similarity 97.1%; Pred. No. 5.3e-124;  
 Matches 640; Conservative 0; Mismatches 9; Indels 10; Gaps 6;  
 QY 612 GAAACACCAAGGCTATGACACAAAGTCGGACATCTGGTCACTGGCATGCAATTTGTATGA 671  
 Db 1 GAAACACCAAGGCTATGACACAAAGTCGGACATCTGGTCACTGGCATGCAATTTGTATGA 60  
 QY 672 GATGTGCTCAGTAATCATGCAATTCGCTGGCTCCAAATTCCTATCCATTGTTTAAAT 731  
 Db 61 GATGTGCTCAGTAATCATGCAATTCGCTGGCTCCAAATTCCTATCCATTGTTTAAAT 120  
 QY 732 TGTGTAAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTAAATGCCATCAT 791

Db 121 TGTGAAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTAAATGCCATCAT 180  
 QY 792 GGAAGCATGTTGAACCAAGAAATCCTTCAATTAAGACCACTCTGCTATCGAAATTTTAAAAAT 851  
 Db 181 GGAAGCATGTTGAACCAAGAAATCCTTCAATTAAGACCACTCTGCTATCGAAATTTTAAAAAT 240  
 QY 852 CCCTTACCTTGTATGAGCAGCTACAGAACCTTAATGTGTA-GATATTAGAAATGACTCTGG 910  
 Db 241 CCCTTACCTTGTATGAGCAGCTACAGAACCTTAATGTGTACGATATTCAGAAATGACTCTGG 300  
 QY 911 AAGACAAAATTTGGATTGTGACAGAGGAGCTGCTCATATAATTAATGCCATGCAAAAAA 970  
 Db 301 AAGACAAAATTTGGATTGTGACAGAGGAGCTGCTCATATAATTAATGCCATGCAAAAAA 360  
 QY 971 GGATCCACCTGACACTCTCAGGGCACTGTGAGAAGTACAGAAAAATGACCCCAAGAGAAA 1030  
 Db 361 GGATCCACCTGACACTCTCAGGGCACTGTGAGAAGTACAGAAAAATGACCCCAAGAGAAA 420  
 QY 1031 GGATCGCGCTGAGGAAGCTCCAGCGGCTCATGAGAAGCCAGGAAGCTGAA-AAAGATT 1089  
 Db 421 GGATCGCGCTGAGGAAGCTCCAGCGGCTCATGAGAAGCCAGGAAGCTGAA-AAAGATT 480  
 QY 1090 GTGGAAGAAAATATGAAGAAAATAGCAAAACGAATGCAAGAA-TTGAGATCTCGGAACTT 1148  
 Db 481 GTGGAAGAAAATATGAAGAAAATAGCAAAACGAATGCAAGAACTTGAGATCTCGGAACTT 540  
 QY 1149 TCAGCAGCTGAGTGTGATGCTACTTCCATGAAAAACACATTT---AAAGGATGG--AA 1203  
 Db 541 TCAGCAGCTGAGTGTGATGCTACTTCCATGAAAAACACATTTTACCAAGGAATGACAG 600  
 QY 1204 GAAAAGGAGGAGCAACCTGGA--GGGAAGACTTTCTTGTTCACCCAGGACGAGATGAA 1260  
 Db 601 AAAAAGGAGGAGCAACCTGGAAGGAGAGACTTTTCTTGTTCACCCAGGACGAGATGAA 659

RESULT 12  
 BI544364 767 bp mRNA linear EST 05-SEP-2001  
 LOCUS 603241760F1 NIH\_MGC\_95 Homo sapiens cDNA clone IMAGE:5284263 5',  
 mRNA sequence.  
 ACCESSION BI544364  
 VERSION BI544364.1 GI:15431676  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE NIH-MGC http://mgs.nci.nih.gov/.  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished (1999)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki  
 Toshiyuki and Piero Carninci (RIKEN)  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM1718 row: g column: 16  
 High quality sequence stop: 728.  
 Location/Qualifiers  
 1..767  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5284263"  
 /tissue\_type="hippocampus"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH\_MGC\_95"  
 /note="Organ: brain; Vector: pBluescriptR (modified

FEATURES  
 source

pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI  
5'-GTGAGTTCCTGGTTTGATTTT-3', size-selected for average  
insert size 2.5 kb and normalized to Ron S. This is a  
primary library enriched for full-length clones and  
constructed using the Cap-trapper method (Carninci, in  
preparation). Library constructed by M. Brownstein  
(NIMH/NHGRI, National Institutes of Health). Note: this  
is a NIH MGCL Library."

## ORIGIN

Query Match	35.2%	Score 510.6	DB 12	Length 767
Best Local Similarity	98.6%	Pred. 4.2e-111		
Matches 547	Conservative 0	Mismatches 4	Indels 4	Gaps 3
QY	1	ATGCTGAAATCCCAAGAGCGAGCTAAGTGTGTGAGTGGATCAACAGCCCAATTTCCACTTAT	60	
Db	216	ATGCTGAAATCCCAAGAGCGAGCTAAGTGTGTGAGTGGATCAACAGCCCAATTTCCACTTAT	275	
QY	61	CAAAAGACCTTGATTGCAAGAAGATACGTGTCTTCAACAAAACTTGGCAGTGGAAAGTTTT	120	
Db	276	CAAAAGACCTTGATTGCAAGAAGATACGTCTTCAACAAAACTTGGCAGTGGAAAGTTTT	335	
QY	121	GGAACTGTCTATCTGGTTTCAGACAGAAAGCCAAACGAGGAGAGGAATTAAGCTACTT	180	
Db	336	GGAACTGTCTATCTGGTTTCAGACAGAAAGCCAAACGAGGAGAGGAATTAAGGTACTT	395	
QY	181	AAGGAAATATCTGTTGGAGAACTAAATCCAAATGAAACTCTACAGGCCCAATTTGGAAGCC	240	
Db	396	AAGGAAATATCTGTTGGAGAACTAAATCCAAATGAAACTCTACAGGCCCAATTTGGAAGCC	455	
QY	241	CAACTCTCTCCAGCTGGACCAACCCAGCCATGTCTAAGTTCATCCAGTTTGTGGAG	300	
Db	456	CAACTCTCTCCAGCTGGACCAACCCAGCCATGTCTAAGTTCATCCAGTTTGTGGAG	515	
QY	301	CAAGATAATTTCTGCATTATCACGAGTACTGTGAGGGCCGAGATCTGGACGATAAAAT	360	
Db	516	CAAGATAATTTCTGCATTATCACGAGTACTGTGAGGGCCGAGATCTGGACGATAAAAT	575	
QY	361	CAGGAATATAACACAGCTGGAAAAATCTTTCCAGAAATCAAAATATAGAACTGGTTTATC	420	
Db	576	CAGGAATATAACACAGCTGGAAAAATCTTTCCAGAAATCAAAATATAGAACTGGTTTATC	635	
QY	421	CAGCTGCTGCTGGGAGTTGAC-TACATGCATGAGAGGAGGATACTTCATCGAGACTTAA	479	
Db	636	CAGCTGCTGCTGGGAGTTGACTTACATGCATGAGAGGAGGATACTTCATCGAGACTTAA	694	
QY	480	GTCAAGAAATGTATTCTTGAAAAATAAATCTCTTAAAAATGGAGATTTCGAGTTTCTCG	539	
Db	695	GTCAAGAAATGTATTCTTGAAAAATAAATCTCTTAAAAATGGAGATTT--GGAGTTTCTCG	752	
QY	540	ACTTCTAAATGGGATC	554	
Db	753	ACTTCTAAATGGGATC	767	

DEPT. T 13

REF ID: A630113  
B11030539

BUSSINESS  
LOCATIONS

NOT  
RECEIVED

**DEFINITION**

SECRET

ACCESSION  
NUMBER

VERSION  
VERSION

KEYWORDS  
corruption

**SOURCE**

## ORGANISMS

11

## REFERENCE

## AUTHORS

ETLII

JOURNAL

COMMENT

Email: cga@bbs-research.nih.gov  
 Tissue Procurement: CLONTECH  
 cDNA Library Preparation: CLONTECH  
 cDNA Library Arrayed by: The  
 DNA Sequencing by: Agencourt  
 Clone distribution: MGC clone  
 found through the I.M.A.G.E.B.  
<http://image.llnl.gov>  
 Plate: LILCM2342 row: o col:  
 High quality sequence stop:

FEATURES source

```

source
1. .867
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6668583"
/lab_host="DH10B (Pl pho
/clone_lib="NH_MGC_83"
/notes="Organ: prostate
Site 1: Sfil (ggccgcctcg
ggccattatggcc); 5' and
as follows: 5' adaptor s
and 3' adaptor sequence:
5'-ATTCTAGAGCCCGAGCGGCG
C, or G and N = A, C, G
kb (range 0.5-4.0 kb) .
by PCR. This library was
and was constructed by C
CAL."

```

## ORIGIN

Query Match	35.2%	Score	509.6	DB	13	Length	867
Best Local Similarity	98.0%	Pred.	No. 7.4e-111				
Matches	537	Conservative	0	Mismatches	9	Indels	2
QY	331	TGTGAGGGCCGAGATCTGGACGATAAAATTCAGGAATATAAACAAGCTGGAATAATCTTT	390				
Db	38	TTTCAGGGCCGAGATCTGGACGATAAAATTCAGGAATATAAACAAGCTGGAATAATCTTT	97				
QY	391	CCAGAAAATCAATAATAGAAATGGTTTATCCAGCTGCTGCTGGAGTTGACTACATGCAT	450				
Db	98	CCAGAAAATCAATAATAGAAATGGTTTATCCAGCTGCTGCTGGAGTTGACTACATGCAT	157				
QY	451	GAGAGGAGGATACCTTCATCGAGAGCTTAAAGTCAAAGAATGTATTCTGAAAAAATAATCTC	510				
Db	158	GAGAGGAGGATACCTTCATCGAGAGCTTAAAGTCAAAGAATGTATTCTGAAAAAATAATCTC	217				
QY	511	CTTAAAAATTCGAGATTTTGGAGTTTCTGCACTTCTAATGGATCCTCTGACCTGGCCACA	570				
Db	218	CTTAAAAATTCGAGATTTTGGAGTTTCTGCACTTCTAATGGATCCTCTGACCTGGCCACA	277				
QY	571	ACTTTAACTCGAACTCCCCATTAATGAGTCTCTGAGGCTCTGAAACACCAAGGCTATGAC	630				
Db	278	ACTTTAACTCGAACTCCCCATTAATGAGTCTCTGAGGCTCTGAAACACCAAGGCTATGAC	337				
QY	631	ACAAAGTCGACATCTGGTCACTGGCAATGCAATTTGTATGAGATGTCTGCAATCAATCAT	690				
Db	338	ACAAAGTCGACATCTGGTCACTGGCAATGCAATTTGTATGAGATGTCTGCAATCAATCAT	397				
QY	691	GCATTGCTGGCTCCAAATTTCTTATCCATTGTTTTTAAAAATTTGTTGAAGGTGCACACCT	750				
Db	398	GCATTGCTGGCTCCAAATTTCTTATCCATTGTTTTTAAAAATTTGTTGAAGGTGCACACCT	457				
QY	751	TCTCTCCCTCGAGAGATATCCAAAAGAACTAAATGCCATCATCGAAAGCATGTTGNACAG	810				
Db	458	TCTCTCCCTCGAGAGATATCCAAACAGAACTAAATGCCATCATCGAAAGCATGTTGNACAG	517				
QY	811	AATCC-TTCATTAGACCATCTGGTATCGAAATTTTAAAAAT-CCCTTACCTTGTATGAGC	868				
Db	518	AATCCTTTTATTAGACCACTGCTATCGAAATTTTAAACATCCCTTACCTTGTATGAGC	577				
QY	869	AGCTACAG	876				

```

Db          578 AGCTACAG 585

RESULT 14
CD106568
LOCUS
DEFINITION
AGENCY: NIH-MGC 13979292 NIH MGC 179 Homo sapiens cDNA clone
IMAGE: 30368885 5', mRNA sequence.
ACCESSION
CD106568
VERSION
CD106568.1 GI:30759742
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 845)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM432 row: 9 column: 06
High quality sequence stop: 719.
Location/Qualifiers
1..845
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30368885"
/issue_type="Pituitary"
/lab_host="PH108-Ton A (T1 and T5 phase resistances)"
/clone_lib="NIH_MGC_179"
/note="Organ: Brain; Vector: pCMV-Sport6.1; Site 1: EcoRV
(destructed); Site 2: NotI; Library is oligo-dr primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.1 kb. Library was
constructed by (Invitrogen). Note: this is a NIH_MGC
Library."

ORIGIN
Query Match 34.8%; Score 504; DB 14; Length 845;
Best Local Similarity 100.0%; Pred. No. 1.6e-109;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 896 CAGAATGACTCTGGAAGACAAAATTTGGATTGTCAGAGGAGGCTGCTCATATAATTA 955
Db 1 CAGAATGACTCTGGAAGACAAAATTTGGATTGTCAGAGGAGGCTGCTCATATAATTA 60

QY 956 ATGCCATGCAAAAAGGATCCACCTGCAGACTCTGAGGCGACTGTGAGAAGTACAGAAA 1015
Db 61 ATGCCATGCAAAAAGGATCCACCTGCAGACTCTGAGGCGACTGTGAGAAGTACAGAAA 120

QY 1016 TCAGCCCAAGAGAGGATGCGGCTGAGAGAGCTCCAGGCGGCTGATGAGAAAGCCAGGA 1075
Db 121 TCAGCCCAAGAGAGGATGCGGCTGAGAGAGCTCCAGGCGGCTGATGAGAAAGCCAGGA 180

QY 1076 AGCTGAAAAGATTTGGAAGAAAATATGAAGAAAATAGCAAAACGAATGCAAGATTGA 1135
Db 181 AGCTGAAAAGATTTGGAAGAAAATATGAAGAAAATAGCAAAACGAATGCAAGATTGA 240

QY 1136 GATCTCGGAACCTTCAGGAGCTGAGTGTGATCTTCATGAAAAAACACATTAAAG 1195
Db 241 GATCTCGGAACCTTCAGGAGCTGAGTGTGATCTTCATGAAAAAACACATTAAAG 300

QY 1196 GAAATGGAAGAAAAGGAGGAGCAACTGAGGGAAGACTTTCTTTGTTCCACCCAGGACGAGG 1255

```

301 GAATGGAAGAAAAGGAGGAGCAACTGAGGAGAACTTTCTTTGTTCCACCCAGGACGAGG 360

1256 ATGAAGAGAGGTGCAAGGAGGAGGAGGAAGAGGAATCTGATGAACCAACTTTAGAGAACTGC 1315

361 ATGAAGAGAGGTGCAAGGAGGAGGAGGAGGAATCTGATGAACCAACTTTAGAGAACTGC 420

1316 CTGAGTCTCAGCCCTATTCCCTTCCATGACCTCCACGAACTTGAATCAATTTGATAGGATG 1375

421 CTGAGTCTCAGCCCTATTCCCTTCCATGACCTCCACGAACTTGAATCAATTTGATAGGATG 480

1376 CCACATCTGACCTTTGGATACCATG 1399

481 CCACATCTGACCTTTGGATACCATG 504

RESULT 15

BB552498

LOCUS

DEFINITION

BB552498 RIKEN full-length enriched, 2 days pregnant adult female

oviduct Mus musculus cDNA clone E230038G21 3', mRNA sequence.

ACCESSION

BB552498

VERSION

BB552498.2 GI:16448169

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 635)

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T.,

Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J.,

Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K.,

Ono, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K.,

Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,

Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,

Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.

RIKEN Mouse ESTs (Arakawa, T., et al. 2001)

Unpublished (2001)

TITLE

JOURNAL

COMMENT

On Aug 1, 2000 this sequence version replaced gi:9638864.

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsr.riken.go.jp,

URL: http://genome.gsc.riken.go.jp/

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,

Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,

Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.

and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,

Sugahara, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

non-redundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yanaka, I.,

Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and

Hayashizaki, Y.

Computational Analysis of Full-length Mouse cDNAs Compared with

Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for

further details.

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.

FEATURES  
source

Location/Qualifiers  
1..635  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="E230038G21"  
/sex="female"  
/tissue\_type="oviduct"  
/dev\_stage="2 days pregnant adult"  
/lab\_host="DH10B"  
/clone\_lib="RIKEN full-length enriched, 2 days pregnant  
adult female oviduct"  
/note="Site\_1: SalI; Site\_2: BamHI; cDNA library was  
prepared and sequenced in Mouse Genome Encyclopedia  
Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in  
RIKEN. Division of Experimental Animal Research in Riken  
contributed to prepare mouse tissues. 1st strand cDNA was  
primed with a primer [5',  
GAGAGAGAGCGCGCAACTCGAGTTTCTTTTCTTNN 3'], cDNA was  
prepared by using trehalose thermo-activated reverse  
transcriptase and subsequently enriched for full-length by  
cap-trapper. Second strand cDNA was prepared with the  
primer adapter of sequence [5',  
GAGAGAGATCTCGAGTTAATAATTATCCCCCCCCCC 3']. cDNA  
was cleaved with BamHI and XhoI. Vector: a modified  
pbluescript KS(+) after bulk excision from lambda FLC I."

ORIGIN

Query Match 32.1%; Score 465.2; DB 10; Length 635;  
Best Local Similarity 83.5%; Pred. No. 2.9e-100;  
Matches 527; Conservative 0; Mismatches 104; Indels 0; Gaps 0;  
QY 460 ATACTTCATCGAGACTTAAGTCAAGATGATTTCTGAAAAATAATCTCCTTAATTT 519  
Db 2 ATACTTCATCGAGACTTGAATCAAGATNATATTTCTGAAAAATAATCTCCTCAAAATC 61  
QY 520 GGAGATTTTGGAGTTTCTCGACTCTTAATGGGATCTGTGACCTGGCCACAACCTTTAACT 579  
Db 62 GGGGATTTTGGAGTTTCTGGCTGTGTAATGGGTTCATGTGAGCTGGCTACACTCTAACC 121  
QY 580 GGAATCTCCCATATATAGTCTGAGGCTCTGAACACCAAGGGCTATGACACAAAGTCG 639  
Db 122 GGGACCCCCANTATATAGTCCGAGGCCCTGAAGCACCAAGGGCTATGATGCCAAGTCT 181  
QY 640 GACATCTGCTCACTGGCATGCAATTTGTATGAGTGTGCTGATGAATCATGCAATTCGCT 699  
Db 182 GACATCTGCTCACTGGCATGCAATTTGTATGAGTGTGCTGATGAATCATGCAATTCGCT 241  
QY 700 GGCTCCAAATTTCTATCCATTTGTTTAAATAATTTGAAGGTGACACACCTTCTCTCCCT 759  
Db 242 GGCTCCAGTTTCTTGCTGTGGTTTGAATATTTGTAAGGTAAACACCTTCACTCCCG 301  
QY 760 GAGAGATATCCAAAGAACTAATGCAATCGATGGAAGGATGTTGAACAGGATCCTTCA 819  
Db 302 GACAGATATCCAGAGAACTAAACACCATGGAACGCAATGTTGAACAGGATCCTTCA 361  
QY 820 TTAAGACCATCTGCTATCGAAATTTTAAATAATCCCTTACCTTGATGAGCAGGTTACAGAAC 879  
Db 362 TTGAGACCGTCGCTGCAGACATTTTAAAGCCCTTACATGAGAGCAGCTTCAGCTC 421  
QY 880 CTAATGTGTAGATATTAGAAATGACTCTGGAAGAGACAAAATTTGGATTGTGAGAGGAG 939  
Db 422 CTGATGTGTAAATACCCAGAGATGACACTGGAAGACAAAGAACTCAGTTTGTGAGAGGAG 481  
QY 940 GCTGCTCATATAATTAATGCCATGCAAAAAGGATCCACCTGCAGACTCTGAGGGCACTG 999  
Db 482 GCTGCTCATACGATTAATGCGGTGCAGAAAAGCTTCACTGCGAGCTCTGCAAGCCTTG 541  
QY 1000 TCAGAGTACAGAAATGACGCCAAGAGAAAGGATCGGCTGAGGAAGCTCCAGGCGGCT 1059

Db 542 TCTGACACGCAGAAAAACGACTCCACAGAACGGATGTGGCTGAGGAAGCTGCAGGAGCC 601  
QY 1060 GATGAGAAAGCCAGGAAGCTGAAAAAGATTG 1090  
Db 602 GATGAGAGAGCCAGGAGGCTGAAAAAGATTG 632

Search completed: October 3, 2004, 00:56:25  
Job time : 4050 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2004, 23:49:06 ; Search time 740 Seconds  
(without alignments)  
9926.760 Million cell updates/sec

Title: US-10-803-278-3

Perfect score: 1449

Sequence: 1 atgttgaaattccaagagggc.....actggaaaaatgaaaaataa 1449

Scoring table: IDENTITY\_NUC

Gapop 10\_0 , Gapext 1.0

Searched: 3340653 seqs, 2334783454 residues

Total number of hits satisfying chosen parameters: 6691306

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US05\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1449	100.0	1449	17	US-10-803-278-3
2	1399	96.5	1938	17	Sequence 3, Appli
3	1399	96.5	2552	13	Sequence 1, Appli
4	1399	96.5	2669	13	Sequence 25, Appli
5	1399	96.5	2801	13	Sequence 2275, A
6	1399	96.5	2847	17	Sequence 18976, A
7	1397.4	96.4	2869	9	Sequence 18976, A
8	1397.4	96.4	2869	17	Sequence 1, Appli
9	1385.4	95.6	2765	17	Sequence 1, Appli
10	759	52.4	2484	13	Sequence 50, Appli
11	669.6	46.2	2538	10	Sequence 16420, A
12	669.6	46.2	2538	10	Sequence 511, App
13	415.8	28.7	451	11	Sequence 511, App
14	395.4	27.3	430	9	Sequence 60, Appli

Query Match 100.0%; Score 1449; DB 17; Length 1449;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1449; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTGAAATTCGAAGGCGAGCTAAGTGTGTGAGTGATCAACAGGCATTTCACCTTAT 60

Db 1 ATGCTGAAATTCGAAGGCGAGCTAAGTGTGTGAGTGATCAACAGGCATTTCACCTTAT 60

QY 61 CCAAGACCTTGATTCGAAGGAGATACGTGCTTCAACAAAACCTTGCAGTGGAGTTT 120

Db 61 CCAAGACCTTGATTCGAAGGAGATACGTGCTTCAACAAAACCTTGCAGTGGAGTTT 120

QY 121 GGAAGTGTCTATCTGTTTCAGACAAAGAACCAACGAGGAGGAATTAAGGTACTT 180

#### ALIGNMENTS

##### RESULT 1

US-10-803-278-3

; Sequence 3, Application US/10803278

; Publication No. US20040175749A1

; GENERAL INFORMATION:

; APPLICANT: Walke, D. Wade

; APPLICANT: Scoville, John

; APPLICANT: Friddle, Carl Johan

; TITLE OF INVENTION: Novel Human Kinases and Polynucleotides Encoding the Same

; FILE REFERENCE: LEX-0348-USA

; CURRENT APPLICATION NUMBER: US/10/803,278

; CURRENT FILING DATE: 2004-03-18

; PRIOR APPLICATION NUMBER: US/10/196,927

; PRIOR FILING DATE: 2002-05-20

; PRIOR APPLICATION NUMBER: US 60/293,248

; PRIOR FILING DATE: 2001-05-24

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 1449

; TYPE: DNA

; ORGANISM: homo sapiens

; US-10-803-278-3

Db 121 GGAAGTCTCTATCTGGTTTCAGACAGAAAGCCAAACGAGGAGGAAATTAAGGTTACTT 180  
 Qy 181 AAGGAAATATCTGTGTGAGAACTAAATCCAAATGAAATCTACAGGCAATTTGGAGCC 240  
 Db 181 AAGGAAATATCTGTGTGAGAACTAAATCCAAATGAAATCTACAGGCAATTTGGAGCC 240  
 Qy 241 CAATCTCTCCAGCTGAGCAACCCAGCCATTTGCAAGTTCCATGCAAGTTTGTGGAG 300  
 Db 241 CAATCTCTCCAGCTGAGCAACCCAGCCATTTGCAAGTTCCATGCAAGTTTGTGGAG 300  
 Qy 301 CAAGATAATTTCTGCAATATACGGAGTACTGTGAGGGCCGAGATCTGGACGATAAAAT 360  
 Db 301 CAAGATAATTTCTGCAATATACGGAGTACTGTGAGGGCCGAGATCTGGACGATAAAAT 360  
 Qy 361 CAGGATATATAAAGCTGGAATAATCTTTCCAGAAAAATCAAAATATAGAAATGTTATC 420  
 Db 361 CAGGATATATAAAGCTGGAATAATCTTTCCAGAAAAATCAAAATATAGAAATGTTATC 420  
 Qy 421 CAGCTGCTGGAGTGTGACTACATGATGAGGAGGATATCTTCATCGAGACTTAAG 480  
 Db 421 CAGCTGCTGGAGTGTGACTACATGATGAGGAGGATATCTTCATCGAGACTTAAG 480  
 Qy 481 TCAAGAAATCTATTTCTGAAAAATAATCTCTTAAATTTGGAGTTTGGAGTTCTCGA 540  
 Db 481 TCAAGAAATCTATTTCTGAAAAATAATCTCTTAAATTTGGAGTTTGGAGTTCTCGA 540  
 Qy 541 CTCTAATGGATCTGTGACCTGGCCACCACTTTAACTGGAACTCCCATATATAGT 600  
 Db 541 CTCTAATGGATCTGTGACCTGGCCACCACTTTAACTGGAACTCCCATATATAGT 600  
 Qy 601 CCTGAGCTCTGAAACACCAAGCTATGACACAAAGTCGAACTCTGGTCACTGGCATGC 660  
 Db 601 CCTGAGCTCTGAAACACCAAGCTATGACACAAAGTCGAACTCTGGTCACTGGCATGC 660  
 Qy 661 ATTTGTATGAGATGTGTCATGATCATGATTCGCTGGCTGCAATTTCTTATCCAT 720  
 Db 661 ATTTGTATGAGATGTGTCATGATCATGATTCGCTGGCTGCAATTTCTTATCCAT 720  
 Qy 721 GTTTTAAATTTCTGAAGTGACACACCTTCTCTCCCTGAGAGATATCAAAAGACTA 780  
 Db 721 GTTTTAAATTTCTGAAGTGACACACCTTCTCTCCCTGAGAGATATCAAAAGACTA 780  
 Qy 781 AATGCCATCATGAAAGCATGTTGAACAAAGATCTTCAATTAAGACCATCTGCTATCGAA 840  
 Db 781 AATGCCATCATGAAAGCATGTTGAACAAAGATCTTCAATTAAGACCATCTGCTATCGAA 840  
 Qy 841 ATTTTAAATTCCTTACCTTGATGAGCACTACAGAACTTAATGTGATGATTTTCAGAA 900  
 Db 841 ATTTTAAATTCCTTACCTTGATGAGCACTACAGAACTTAATGTGATGATTTTCAGAA 900  
 Qy 901 ATGACTCTGGAAGACAAAAATTTGATTTGATGTCAGAGGGCTGCTCATATATTAATGCC 960  
 Db 901 ATGACTCTGGAAGACAAAAATTTGATTTGATGTCAGAGGGCTGCTCATATATTAATGCC 960  
 Qy 961 ATGCANAAAGATCCACCTGAGACTCTGAGGCACTGTGAGAGTACAGAAATGACG 1020  
 Db 961 ATGCANAAAGATCCACCTGAGACTCTGAGGCACTGTGAGAGTACAGAAATGACG 1020  
 Qy 1021 CCAAGAAAGAGGATCGGCTGAGGAGCTCCAGGGGGCTGATGAGAAAGCCAGGAGCTG 1080  
 Db 1021 CCAAGAAAGAGGATCGGCTGAGGAGCTCCAGGGGGCTGATGAGAAAGCCAGGAGCTG 1080  
 Qy 1081 AAAAGATTTGGAAGAAAAATATGAAGAAATAGCAAAAGATTCGAAGATTTGATGATCT 1140  
 Db 1081 AAAAGATTTGGAAGAAAAATATGAAGAAATAGCAAAAGATTCGAAGATTTGATGATCT 1140  
 Qy 1141 CGGAATTTTCAGAGCTGAGTGTGATGATCTTCCATGAAAAACACATTTAAAGGAATG 1200  
 Db 1141 CGGAATTTTCAGAGCTGAGTGTGATGATCTTCCATGAAAAACACATTTAAAGGAATG 1200  
 Qy 1201 GAAGAAAGGAGGAGCACTGAGGAGAGACTTTCTTGTTCACCCAGGACGAGATGAA 1260

Db 1201 GAAGAAAGGAGGAGCAACCTGAGGAGAGACTTTCTTGTTCACCCAGGAGGAGATGAA 1260  
 Qy 1261 GAGAGTGGCAAGGAGGAGAGAGAAATCTGATGAACCAACTTTAGAGAACTCTGCTGAG 1320  
 Db 1261 GAGAGTGGCAAGGAGGAGAGAGAAATCTGATGAACCAACTTTAGAGAACTCTGCTGAG 1320  
 Qy 1321 TCTGAGCTATCTCTTCCATGAGACTCTCAAGAACTTGAATCAATTTAGAGGATGCGACA 1380  
 Db 1321 TCTGAGCTATCTCTTCCATGAGACTCTCAAGAACTTGAATCAATTTAGAGGATGCGACA 1380  
 Qy 1381 TCTGAGCTTGTATACCATGAGACTGTAAATCTAATTTCACTAGAGAAATGCGAAAAAT 1440  
 Db 1381 TCTGAGCTTGTATACCATGAGACTGTAAATCTAATTTCACTAGAGAAATGCGAAAAAT 1440  
 Qy 1441 GAAAAATAA 1449  
 Db 1441 GAAAAATAA 1449

RESULT 2  
 US-10-803-278-1  
 ; Sequence 1, Application US/10803278  
 ; Publication No. US2004017549A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Walke, D. Wade  
 ; APPLICANT: Scoville, John  
 ; APPLICANT: Friddle, Carl Johan  
 ; TITLE OF INVENTION: Novel Human Kinases and Polynucleotides Encoding the Same  
 ; FILE REFERENCE: LEX-0348-USA  
 ; CURRENT APPLICATION NUMBER: US/10/803,278  
 ; CURRENT FILING DATE: 2004-03-18  
 ; PRIOR APPLICATION NUMBER: US/10/196,927  
 ; PRIOR FILING DATE: 2002-05-20  
 ; PRIOR APPLICATION NUMBER: US 60/293,248  
 ; PRIOR FILING DATE: 2001-05-24  
 ; NUMBER OF SEQ ID NOS: 5  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1  
 ; LENGTH: 1938  
 ; TYPE: DNA  
 ; ORGANISM: homo sapiens  
 US-10-803-278-1

Query Match 96.5%; Score 1399; DB 17; Length 1938;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCTGAAATTCAGAGGCGAGTAAAGTGTGAGTGGATCAACAGCCATTTCCACTTAT 60  
 Db 1 ATGCTGAAATTCAGAGGCGAGTAAAGTGTGAGTGGATCAACAGCCATTTCCACTTAT 60  
 Qy 61 CCAAGAACCTTGTATGCAAGAGATACGCTCTTCAACAAAACTTGGCAGTGGAGTTT 120  
 Db 61 CCAAGAACCTTGTATGCAAGAGATACGCTCTTCAACAAAACTTGGCAGTGGAGTTT 120  
 Qy 121 GGAATCTGTATCTGTTTCAGACAGAAAGCCAAACGAGGAGGAGAAATTAAGGTTACT 180  
 Db 121 GGAATCTGTATCTGTTTCAGACAGAAAGCCAAACGAGGAGGAGAAATTAAGGTTACT 180  
 Qy 181 AAGGAATATCTGTTGGAGAACTAAATCCAAATGAAATCTGACAGCCCAATTTGGAAGCC 240  
 Db 181 AAGGAATATCTGTTGGAGAACTAAATCCAAATGAAATCTGACAGCCCAATTTGGAAGCC 240  
 Qy 241 CAATCTCTCCAGCTGAGCAACCCAGCCATTTGCAAGTTCCATGCAAGTTTGTGGAG 300  
 Db 241 CAATCTCTCCAGCTGAGCAACCCAGCCATTTGCAAGTTCCATGCAAGTTTGTGGAG 300  
 Qy 301 CAAGATAATTTCTGCAATATACGGAGTACTGTGAGGGCCGAGATCTGGACGATAAAAT 360  
 Db 301 CAAGATAATTTCTGCAATATACGGAGTACTGTGAGGGCCGAGATCTGGACGATAAAAT 360  
 Qy 361 CAGGATATATAAAGCTGGAATAATCTTTCCAGAAAAATCAAAATATAGAAATGTTATC 420

Db 361 CAGGAATATAAACAGCTGGAAAAATCTTTCCAGAAAAATCAATAATAGATGGTTTATC 420  
QY 421 CAGCTCTCTGCTGGAGTTGACTACATGATGAGGAGGAGGATCTTCATCGAGACTTAAG 480  
Db 421 CAGCTCTCTGCTGGAGTTGACTACATGATGAGGAGGAGGATCTTCATCGAGACTTAAG 480  
QY 481 TCAAGAAATGTATTTCTGAAAAATATCTCTTAAATTTGGAGATTTTGGAGTTTCTCGA 540  
Db 481 TCAAGAAATGTATTTCTGAAAAATATCTCTTAAATTTGGAGATTTTGGAGTTTCTCGA 540  
QY 541 CTTCTAATGGGATCCTGTGACCTGGCCACAACTTTAACTGGAACTCCCATATATGAT 600  
Db 541 CTTCTAATGGGATCCTGTGACCTGGCCACAACTTTAACTGGAACTCCCATATATGAT 600  
QY 601 CTTGAGGCTCTGAAACACCAAGCTATGACACAAAGTCGGACATCTGTCTCACTGGCATGC 660  
Db 601 CTTGAGGCTCTGAAACACCAAGCTATGACACAAAGTCGGACATCTGTCTCACTGGCATGC 660  
QY 661 ATTGTTGATGAGATGTGCTGCATGATCATGATTCGCTGGCTCCAAATTTCTTATCCAT 720  
Db 661 ATTGTTGATGAGATGTGCTGCATGATCATGATTCGCTGGCTCCAAATTTCTTATCCAT 720  
QY 721 GTTTTAAAAATTTTCAAGGTGACACACCTTCTCTCCCTGAGAGATATCAAAAAGACTA 780  
Db 721 GTTTTAAAAATTTTCAAGGTGACACACCTTCTCTCCCTGAGAGATATCAAAAAGACTA 780  
QY 781 ATGCCATCATGGAAGAGCTTTGAACAAGATCTCTTCAATTAAGACCATCTCTGCTATCGAA 840  
Db 781 ATGCCATCATGGAAGAGCTTTGAACAAGATCTCTTCAATTAAGACCATCTCTGCTATCGAA 840  
QY 841 ATTTTAAAAATCCCTTACCTTGTATGAGCAGCTACAGAACCTAATGTGTAGATATTCAGAA 900  
Db 841 ATTTTAAAAATCCCTTACCTTGTATGAGCAGCTACAGAACCTAATGTGTAGATATTCAGAA 900  
QY 901 ATGACTCTCGAAGACAAAAATTTGGATTTGTCTGAGAGGAGCTGCTCATATAATTAATGCC 960  
Db 901 ATGACTCTCGAAGACAAAAATTTGGATTTGTCTGAGAGGAGCTGCTCATATAATTAATGCC 960  
QY 961 ATGCAAAAAGAGTCCACTGAGACTCTGAGGGAAGCTCTGAGGAGCTGATGAGAAAGCTG 1020  
Db 961 ATGCAAAAAGAGTCCACTGAGACTCTGAGGGAAGCTCTGAGGAGCTGATGAGAAAGCTG 1020  
QY 1021 CCAAGAGAAAGGATGGCTGAGGAAAGCTCCAGGCGGCTGATGAGAAAGCTGAGGAGGCTG 1080  
Db 1021 CCAAGAGAAAGGATGGCTGAGGAAAGCTCCAGGCGGCTGATGAGAAAGCTGAGGAGGCTG 1080  
QY 1081 AAAAGATTTGGAAGAAAAATATGAAGAAATAGCAAAAGAAATGCAAGATGCAAGATTTGATCT 1140  
Db 1081 AAAAGATTTGGAAGAAAAATATGAAGAAATAGCAAAAGAAATGCAAGATGCAAGATTTGATCT 1140  
QY 1141 CGGAACCTTTGAGAGCTGAGTGTGATGATCTCCATGAAAAACACATTTAAAGGAATG 1200  
Db 1141 CGGAACCTTTGAGAGCTGAGTGTGATGATCTCCATGAAAAACACATTTAAAGGAATG 1200  
QY 1201 GAAGAAAGGAGGAGCAACTGAGGGAAGACTTTTGTGTTCACTCCAGGAGGAGGATGAA 1260  
Db 1201 GAAGAAAGGAGGAGCAACTGAGGGAAGACTTTTGTGTTCACTCCAGGAGGAGGATGAA 1260  
QY 1261 GAGAGTGGCAGGCGAGGAGGAGGATCTGATGACCAACTTTAGAGAACTGCTGCTGAG 1320  
Db 1261 GAGAGTGGCAGGCGAGGAGGAGGATCTGATGACCAACTTTAGAGAACTGCTGCTGAG 1320  
QY 1321 TCTCAGCTATTCTCTCCATGACCTCCAGCACTTGAATCAATTTAGAGGATGCCACA 1380  
Db 1321 TCTCAGCTATTCTCTCCATGACCTCCAGCACTTGAATCAATTTAGAGGATGCCACA 1380  
QY 1381 TCTGACCTTGGATPACCATG 1399  
Db 1381 TCTGACCTTGGATPACCATG 1399

RESULT 3

US-10-333-314-25

; Sequence 25, Application US/10333314  
; Publication No. US20030211093A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.; YUE, Henry  
; APPLICANT: KHAN, Farrah A.; GURURAJAN, Rajagopal  
; APPLICANT: HAPALIA, April J.A.; CHAWLA, Jayalaxmi  
; APPLICANT: ARVIZU, Chandra S.; RAMKUMAR, Jennifer L.  
; APPLICANT: GAUNDHI, Ameena R.; POLICKY, Jennifer L.  
; APPLICANT: BAUGHN, Mariah R.; TRIBOULEY, Catherine M.  
; APPLICANT: THORNTON, Michael B.; BANDMAN, Olga  
; APPLICANT: NGUYEN, Dannie B.; LU, Yan  
; APPLICANT: BURFORD, Neil; LAL, Preeti G.  
; APPLICANT: DING, Li; YAO, Monique G.  
; APPLICANT: ELLIOTT, Vicki S.; RECIPON, Shirley A.  
; APPLICANT: KEARNEY, Liam; LU, Dying Aina M.  
; APPLICANT: GREENWALD, Sara R.; TANG, Y. Tom  
; APPLICANT: XU, Yiming; WALSH, Roderick T.  
; APPLICANT: GISTZEN, Kimberly J.; YANG, Junming  
; APPLICANT: JACKSON, Jennifer L.  
; TITLE OF INVENTION: HUMAN KINASES  
; FILE REFERENCE: PI-0162 USN  
; CURRENT APPLICATION NUMBER: US/10/333,314  
; PRIOR FILING DATE: 2003-01-15  
; PRIOR APPLICATION NUMBER: PCT/US01/23092  
; PRIOR FILING DATE: 2001-07-20  
; PRIOR APPLICATION NUMBER: US 60/220,038  
; PRIOR FILING DATE: 2000-07-21  
; PRIOR APPLICATION NUMBER: US 60/222,112  
; PRIOR FILING DATE: 2000-07-28  
; PRIOR APPLICATION NUMBER: US 60/222,831  
; PRIOR FILING DATE: 2000-08-04  
; PRIOR APPLICATION NUMBER: US 60/224,729  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: PERL Program  
; SEQ ID NO 25  
; LENGTH: 2552  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No: 7477585CB1  
US-10-333-314-25

Query Match 96.5%; Score 1399; DB 13; Length 2552;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGCTGAAATTCGAAGAGGAGCTAAGTGTGTGAGTGGATCAACAGCCATTTCCACTTAT 60  
Db 278 ATGCTGAAATTCGAAGAGGAGCTAAGTGTGTGAGTGGATCAACAGCCATTTCCACTTAT 337  
QY 61 CCAAGACCTTGATGCAAGAGATAGTGTCTTCAACAAAACTTGGCAGTGGAGTTT 120  
Db 338 CCAAGACCTTGATGCAAGAGATAGTGTCTTCAACAAAACTTGGCAGTGGAGTTT 397  
QY 121 GGAAGTGTCTATCTGTTTTCAGACAAAGAAAGCCAAAGAGGAGGAAATTAAGGTA 180  
Db 398 GGAAGTGTCTATCTGTTTTCAGACAAAGAAAGCCAAAGAGGAGGAAATTAAGGTA 457  
QY 181 AAGGAATATCTGTGAGAACTAAATCCAAATGAAAATCTGACGGCCATTTGGAGGCC 240  
Db 458 AAGGAATATCTGTGAGAACTAAATCCAAATGAAAATCTGACGGCCATTTGGAGGCC 517  
QY 241 CAAGTCTCTCAAGCTGGACCCAGCCATTTCAAGTTCCATGCAAGTTTGTGGAG 300  
Db 518 CAAGTCTCTCAAGCTGGACCCAGCCATTTCAAGTTCCATGCAAGTTTGTGGAG 577  
QY 301 CAAGATAATTTCTGCATTATCAAGGAGTACTGTGAGGCGGCGAGATCTGGACGATAAAT 360  
Db 578 CAAGATAATTTCTGCATTATCAAGGAGTACTGTGAGGCGGCGAGATCTGGACGATAAAT 637  
QY 361 CAGGAATATAAACAGCTGGAAAAATCTTTCCAGAAAAATCAATAATAGATGGTTTATC 420

638	Db	CAGGAATATAAA	CAAGCTGGA	AAAAAATTCTTCCAGAAAATCAAATAATAGAAATCGTTTATC	697
421	Qy	CAGCTGCTGTCTGGAGTTGACTCATCATGCATGAGAGGAGTA	CTTCAATCGAGACTTAAAG	480	
698	Db	CAGCTGCTGTGGAGTTGACTCATGCATGAGAGGAGTA	CTTCATCGAGACTTAAAG	757	
481	Qy	TCAAAGAATCTATTTCTGAAAAATAATCTCCTTAAAAATCGAGATTTTGAGTTCTCTCGA	540		
758	Db	TCAAAGAATCTATTTCTGAAAAATAATCTCCTTAAAAATCGAGATTTTGAGTTCTCTCGA	817		
541	Qy	CTTCTAATGGATCCTGTGACCTGGCCACAACTTTAACTCGGAATCCCCCAATATATGAGT	600		
818	Db	CTTCTAATGGATCCTGTGACCTGGCCACAACTTTAACTCGGAATCCCCCAATATATGAGT	877		
601	Qy	CCTGAGGCTCTGAAAACAACCAAGGCTATGACACAAAATCGCAATCTGGTCACTGGCATGC	660		
878	Db	CCTGAGGCTCTGAAAACAACCAAGGCTATGACACAAAATCGCAATCTGGTCACTGGCATGC	937		
561	Qy	ATTTTGTATCAGATGTCTGCATGAATCATGCAATCGCTGGCTCCAATTTCTTTATCCATT	720		
938	Db	ATTTTGTATCAGATGTCTGCATGAATCATGCAATCGCTGGCTCCAATTTCTTTATCCATT	997		
721	Qy	GTTTTAAAAATTTGTTGAAGGTGACACACCCTTCTCTCCCTGAGAGATATCAAAAAGAACTA	780		
998	Db	GTTTTAAAAATTTGTTGAAGGTGACACACCCTTCTCTCCCTGAGAGATATCAAAAAGAACTA	1057		
781	Qy	AATGCCATCATGGAAGCATGTTGAACAAGATCCCTTATTAGAGCATCTGCTATCGAA	840		
1058	Db	AATGCCATCATGGAAGCATGTTGAACAAGATCCCTTATTAGAGCATCTGCTATCGAA	1117		
841	Qy	ATTTTAAAAATCCCCTTACCTTGATGAGCAGCTACAGAACCTTAATGTGTAGATATTCAGAA	900		
1118	Db	ATTTTAAAAATCCCCTTACCTTGATGAGCAGCTACAGAACCTTAATGTGTAGATATTCAGAA	1177		
901	Qy	ATGACTCTGAAGACAAAAATTTGGATGTTCAGNAGGAGCTGCTCATTAATAATATGCC	960		
1178	Db	ATGACTCTGAAGACAAAAATTTGGATGTTCAGNAGGAGCTGCTCATTAATAATATGCC	1237		
961	Qy	ATGCAAAAAAGGATPCCACCTCGACACTCTGAGGCACTGTGAGAGTACAGAAAATGACG	1020		
1238	Db	ATGCAAAAAAGGATPCCACCTCGACACTCTGAGGCACTGTGAGAGTACAGAAAATGACG	1297		
1021	Qy	CAAAGAAAAGGATCGCGCTGAGGAAGCTCGAGCGCTCATGAGAAAGCCAGGAAGCTG	1080		
1298	Db	CAAAGAAAAGGATCGCGCTGAGGAAGCTCGAGCGCTCATGAGAAAGCCAGGAAGCTG	1357		
1081	Qy	AAAAAGATTGTGAAGAAAAAATTAAGAAAAATAGCAAAAGCAATGCAAGAAATTGAGATCT	1140		
1358	Db	AAAAAGATTGTGAAGAAAAAATTAAGAAAAATAGCAAAAGCAATGCAAGAAATTGAGATCT	1417		
1141	Qy	CGGAACTTTTCAGACGTGAGTGTGTATGTATCTCATGAAAAAACAATTTAAAGGAATG	1200		
1418	Db	CGGAACTTTTCAGACGTGAGTGTGTATGTATCTCATGAAAAAACAATTTAAAGGAATG	1477		
1201	Qy	GAAGAAAGGAGGACAACTGAGGGAAGACTTCTTGTTCACCCAGGACGAGGATGAA	1260		
1478	Db	GAAGAAAGGAGGACAACTGAGGGAAGACTTCTTGTTCACCCAGGACGAGGATGAA	1537		
1261	Qy	GAGAGTGGCAAGGACAGGGAAGAGGAATCTGATGAACCAACTTTTAGAGAACCTTCCCTGAG	1320		
1538	Db	GAGAGTGGCAAGGACAGGGAAGAGGAATCTGATGAACCAACTTTTAGAGAACCTTCCCTGAG	1597		
1321	Qy	TCTCAGCCTATTCCCTTCATGGACCTCCACGAACTTCAATCAATTTAGAGGATGCCACA	1380		
1598	Db	TCTCAGCCTATTCCCTTCATGGACCTCCACGAACTTCAATTTAGAGGATGCCACA	1657		
1381	Qy	TCTGACCTTGGATACCATG	1399		
1658	Db	TCTGACCTTGGATACCATG	1676		

```

US-10-425-114-22275
; Sequence 22275, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 22275
; LENGTH: 2669
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3547-006-G1_FLI
; US-10-425-114-22275

```

Query Match	96.5%;	Score 1399;	DB 13;	Length 2669;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1399;	Conservative	0;	Mismatches	0; Indels 0; Gaps

  

Qy	1	ATGCTGAATAATCCAAAGAGCGAGCTAAGTGTGTGAGTCAGATCAACAGCATTTCCTCACTTAT	60
Dd	48	ATGCTGAATAATCCAAAGAGCGAGCTAAGTGTGTGAGTCAGATCAACAGCATTTCCTCACTTAT	107
Qy	61	CAAAGAAGCCTTGATTTGCAAGAAGATAAGTGTCTCAACAATACTTGCGAGTCGAAGTTTT	120
Dd	108	CCAAGAAGCCTTGATTTGCAAGAAGATAAGTGTCTCAACAATACTTGCGAGTCGAAGTTTT	167
Qy	121	GGAACCTGCTACTGCTGTTTCAGACAAGAAAACCACGAGGAGAGAAATTAAGAGTACTT	180
Dd	168	GGAACTGCTACTGCTGTTTCAGACAAGAAAACCACGAGGAGAGAAATTAAGAGTACTT	227
Qy	181	AAGAAAAATATCTGTTGGAGAACTAAATCCAAATGAAACTGTACAGGCCAAATTTGGAAGCC	240
Dd	228	AAGAAAAATATCTGTTGGAGAACTAAATCCAAATGAAACTGTACAGGCCAAATTTGGAAGCC	287
Qy	241	CACCTCCTCTCCAGCTCGACCGACCCAGGCATTTGTCAAGTTCCATCAAGTTTTTGTGGAG	300
Dd	288	CACCTCCTCTCCAGCTCGACCGACCCAGGCATTTGTCAAGTTCCATCAAGTTTTTGTGGAG	347
Qy	301	CAAGATAAATTTCTGCATTATCACGAGTACTGTGAGGCCGAGATCTGACGATCAAAAT	360
Dd	348	CAAGATAAATTTCTGCATTATCACGAGTACTGTGAGGCCGAGATCTGACGATCAAAAT	407
Qy	361	CAGGAATATAACRAGCTGGAAAAATCTTTCGAGAAATCAAAATATAGATGGTTTTATC	420
Dd	408	CAGGAATATAAACRAGCTGGAAAAATCTTTCGAGAAATCAAAATATAGATGGTTTTATC	467
Qy	421	CAGCTGCTGCTGGGAGTTGACTACATCATCAGAGGAGGATCTTCATCGAGACTTAAG	480
Dd	468	CAGCTGCTGCTGGGAGTTGACTACATCATCAGAGGAGGATCTTCATCGAGACTTAAG	527
Qy	481	TCAAGAAATGATATTTCTGAAAAAATAATCTCTTAAATTTGGAGATTTTGGAGTTTCTCGA	540
Dd	528	TCAAGAAATGATATTTCTGAAAAAATAATCTCTTAAATTTGGAGATTTTGGAGTTTCTCGA	587
Qy	541	CTTCTAAATGGGATCTGTGACCTGGCCACAACTTTAACTGGAACTCCCCATTATATGAGT	600
Dd	588	CTTCTAAATGGGATCTGTGACCTGGCCACAACTTTAACTGGAACTCCCCATTATATGAGT	647
Qy	601	CCTGAGGCTCTGAARACCAAGGCTATGACACAAGTCGACATCTGGTCACTGGCATGC	660
Dd	648	CCTGAGGCTCTGAARACCAAGGCTATGACACAAGTCGACATCTGGTCACTGGCATGC	707
Qy	661	ATTTTGTATGAGATGCTGTCATGAATCATGCAATTCGCTGGCTCCAAATTTCTTATCCATT	720



1168 ATGCAAAAAGGATCACTGCTGAGATCTGAGGCACTGTCAGAGTACAGAAAATGACG 1227  
1021 CCAAGAAAGGATGCGGCTGAGGAAGCTCCAGGCGGCTGATGAGAAAGCCAGGAGCTG 1080  
1228 CCAAGAAAGGATGCGGCTGAGGAAGCTCCAGGCGGCTGATGAGAAAGCCAGGAGCTG 1287  
1081 AAAAAAGTTGTGAAGAAATAATGAGAAATAGCAAAAGCAATGCAAGAAATGAGATCT 1140  
1288 AAAAAAGTTGTGAAGAAATAATGAGAAATAGCAAAAGCAATGCAAGAAATGAGATCT 1347  
1141 CGGAATTTTCAGCAGCTGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1200  
1348 CGGAATTTTCAGCAGCTGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 1407  
1201 GAAGAAAGGAGGCAACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260  
1408 GAAGAAAGGAGGCAACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1467  
1261 GAGAGTGGCAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320  
1468 GAGAGTGGCAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1527  
1321 TCTCAGCTATTTCTTCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380  
1528 TCTCAGCTATTTCTTCCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1587  
1381 TCTGACCTTGGATACCATG 1399  
1588 TCTGACCTTGGATACCATG 1606

## RESULT 6

US-10-803-278-5  
; Sequence 5: Application US/10803278  
; Publication NO: US2004017549A1  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Scoville, John  
; APPLICANT: Friddle, Carl Johan  
; TITLE OF INVENTION: Novel Human Kinases and Polynucleotides Encoding the Same  
; FILE REFERENCE: LEX-0348-USA  
; CURRENT APPLICATION NUMBER: US/10/803,278  
; CURRENT FILING DATE: 2004-03-18  
; PRIOR APPLICATION NUMBER: US/10/196,927  
; PRIOR FILING DATE: 2002-05-20  
; PRIOR APPLICATION NUMBER: US 60/293,248  
; PRIOR FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 2847  
; TYPE: DNA  
; ORGANISM: homo sapiens  
US-10-803-278-5

Query Match 96.5%; Score 1399; DB 17; Length 2847;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGCTGAAATCCAGAGGCGAGCTAAGTGTGTGAGTGGATCAACAGCCATTTCACCTTAT 60  
319 ATGCTGAAATCCAGAGGCGAGCTAAGTGTGTGAGTGGATCAACAGCCATTTCACCTTAT 378  
61 CCAAGACCTTGTATGCAAGAGATAGTCTTCAACAAAACCTTGCAGTGGAGGTTTT 120  
379 CCAAGACCTTGTATGCAAGAGATAGTCTTCAACAAAACCTTGCAGTGGAGGTTTT 438  
121 GGAATGTCTATCTGTTTTCAGACAAGAAAGCCAAACGAGGAGGAGGAAATTAAGGTA 180  
439 GGAATGTCTATCTGTTTTCAGACAAGAAAGCCAAACGAGGAGGAGGAAATTAAGGTA 498  
181 AAGGAATATCTGTTGGAGAACTAAATCCAAATGTAAGGAGGAGGAGGAGGAGGAGGAGG 240

499 AAGGAATATCTGTTGGAGAACTAAATCCAAATGAAACTGTACAGGCCAATTTGGAAGCC 558  
241 CAATCTCTCCAGCTGGAGCCACCCAGCCATTTGCAAGTTCCATGCAAGTTTGTGGAG 300  
559 CAATCTCTCCAGCTGGAGCCACCCAGCCATTTGCAAGTTCCATGCAAGTTTGTGGAG 618  
301 CAAGATAATTTCTGCAATTAATCAAGAGTACTGTGAGGCGCGAGATCTGAGACGATAAAT 360  
619 CAAGATAATTTCTGCAATTAATCAAGAGTACTGTGAGGCGCGAGATCTGAGACGATAAAT 678  
361 CAGGAATATAACAGAGCTGGAAAATCTTTCCAGAAAATCAAAATAATAGAAATGTTTATC 420  
679 CAGGAATATAACAGAGCTGGAAAATCTTTCCAGAAAATCAAAATAATAGAAATGTTTATC 738  
421 CAGCTGCTGCTGGAGTTGACTACATGATGAGAGGAGGATCTTCAATCGAGACTTAAAG 480  
739 CAGCTGCTGCTGGAGTTGACTACATGATGAGAGGAGGATCTTCAATCGAGACTTAAAG 798  
481 TCAAGAAATGTAATTTCTGAAAATTAATCTCTTAAANTTGGAGATTTTGGAGTTTCTCGA 540  
799 TCAAGAAATGTAATTTCTGAAAATTAATCTCTTAAANTTGGAGATTTTGGAGTTTCTCGA 858  
541 CTTCTAATGGGATCCTGTGACCTGGCCACAACTTTAACTGAACTCCCAATTTATGAGT 600  
859 CTTCTAATGGGATCCTGTGACCTGGCCACAACTTTAACTGAACTCCCAATTTATGAGT 918  
601 CCGAGGCTCTGAAAACCAAGGCTATGACAAAAGTCGGACATCTGCTCACTGGCATGC 660  
919 CCGAGGCTCTGAAAACCAAGGCTATGACAAAAGTCGGACATCTGCTCACTGGCATGC 978  
661 ATTTTGTATGAGATGCTGTCATGAATCATGCAATTCGCTGGCTCCAAATTTCTTATCCATT 720  
979 ATTTTGTATGAGATGCTGTCATGAATCATGCAATTCGCTGGCTCCAAATTTCTTATCCATT 1038  
721 GTTTTAAAAATTTGTGAAGTGAACACCTTCTCTCCCTGAGAGATATCCAAAAGAACTA 780  
1039 GTTTTAAAAATTTGTGAAGTGAACACCTTCTCTCCCTGAGAGATATCCAAAAGAACTA 1098  
781 AATGCCATCATGGAAGAGCTTTGAAGAAGATCCTTCATTAAGACCATCTGCTATCGAA 840  
1099 AATGCCATCATGGAAGAGCTTTGAAGAAGATCCTTCATTAAGACCATCTGCTATCGAA 1158  
841 ATTTTAAAAATTTGTGAAGTGAACACCTTCTCTCCCTGAGAGATATCCAAAAGAACTA 900  
1159 ATTTTAAAAATTTGTGAAGTGAACACCTTCTCTCCCTGAGAGATATCCAAAAGAACTA 1218  
901 ATGACTCTGGAAGACAAAATTTGATGTCAGAGGAGGCTGCTCATATAATTAATGCC 960  
1219 ATGACTCTGGAAGACAAAATTTGATGTCAGAGGAGGCTGCTCATATAATTAATGCC 1278  
961 ATGCAAAAAAGGATCCACCTGAGACTCTGAGGCGACTGTGAGAAAGTACAGAAATGACG 1020  
1279 ATGCAAAAAAGGATCCACCTGAGACTCTGAGGCGACTGTGAGAAAGTACAGAAATGACG 1338  
1021 CCAAGAAAGGATGCTGGCTGAGGAGCTCCAGGCGGCTGATGAGAAAGCCAGGAGCTG 1080  
1339 CCAAGAAAGGATGCTGGCTGAGGAGCTCCAGGCGGCTGATGAGAAAGCCAGGAGCTG 1398  
1081 AAAAGATTTGTGAAGAAAAATATGAAGAAATAGCAAAAGCAATTTAAAGAGGATG 1140  
1399 AAAAGATTTGTGAAGAAAAATATGAAGAAATAGCAAAAGCAATTTAAAGAGGATG 1458  
1141 CGGAATTTTCAGCAGCTGAGTGTGATGATCTCCATGAAAAACACATTTAAAGAGGATG 1200  
1459 CGGAATTTTCAGCAGCTGAGTGTGATGATCTCCATGAAAAACACATTTAAAGAGGATG 1518  
1201 GAAGAAAGGAGGAGCACTTGGAGAGAGCTTTCTGTTTCAACCCAGGAGGAGGAGGAGG 1260  
1519 GAAGAAAGGAGGAGCACTTGGAGAGAGCTTTCTGTTTCAACCCAGGAGGAGGAGGAGG 1578  
1261 GAGAGTGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320

Db 1579 GAGAGGTGGCAAGGACGGAAGAGGAATCTGATGAACCAACTTTAGAGAACCTGCCTGAG 1638  
Qy 1321 TCTCAGCCTATTCCTCCATGACCTCCAGAACTTGAATCAATTTAGAGGATGCCACA 1380  
Db 1639 TCTCAGCCTATTCCTCCATGACCTCCAGAACTTGAATCAATTTAGAGGATGCCACA 1698  
Qy 1381 TCTGACCTTGGATACCATG 1399  
Db 1699 TCTGACCTTGGATACCATG 1717

## RESULT 7

US-09-731-231A-1  
; Sequence 1, Application US/09731231A  
; Patent No. US20020082189A1  
; GENERAL INFORMATION:  
; APPLICANT: GUEGLER, Karl et al.  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001007  
; CURRENT APPLICATION NUMBER: US/09/731,231A  
; CURRENT FILING DATE: 2000-12-07  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 2869  
; TYPE: DNA  
; ORGANISM: Human  
US-09-731-231A-1

Query Match 96.4%; Score 1397.4; DB 9; Length 2869;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1398; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 ATGCTGAAATTCACAGGAGCAGTAAAGTGTGAGTGGATCAACGCCATTTCCACTTAT 60  
Db 181 ATGCTGAAATTCACAGGAGCAGTAAAGTGTGAGTGGATCAACGCCATTTCCACTTAT 240  
Qy 61 CCAAGACCTTGAATTCAGAGAGATACGTGCTTCAACAAAACCTTGGCAGTGGAAATTT 120  
Db 241 CCAAGACCTTGAATTCAGAGAGATACGTGCTTCAACAAAACCTTGGCAGTGGAAATTT 300  
Qy 121 GGAAGTCTATCTGCTTTTCAGACAGAAAGCCACGAGGAGGAGGAAATTAAGTACTT 180  
Db 301 GGAAGTCTATCTGCTTTTCAGACAGAAAGCCACGAGGAGGAGGAAATTAAGTACTT 360  
Qy 181 AAGGAATATCTGTTGAGAACTTAAATCCAAATGAACTGTACAGGCCAAATTTGGAAGCC 240  
Db 361 AAGGAATATCTGTTGAGAACTTAAATCCAAATGAACTGTACAGGCCAAATTTGGAAGCC 420  
Qy 241 CCACTCTCTCCAGCTGACACCAACCCAGCAGTGTCAAGTTCATGCAAGTTTGTGGAG 300  
Db 421 CCACTCTCTCCAGCTGACACCAACCCAGCAGTGTCAAGTTCATGCAAGTTTGTGGAG 480  
Qy 301 CAAGATAATTTCTGCAATTCACGGAGTACTGTGAGGCGGAGATCTGACGATATAAAT 360  
Db 481 CAAGATAATTTCTGCAATTCACGGAGTACTGTGAGGCGGAGATCTGACGATATAAAT 540  
Qy 361 CAGGAATATAACAGCTGGAAAAATCTTTCCAGAAAAATCAAAATAGAAATGTTTATC 420  
Db 541 CAGGAATATAACAGCTGGAAAAATCTTTCCAGAAAAATCAAAATAGAAATGTTTATC 600  
Qy 421 CAGCTGCTCTGGAGTTCACATGATGATGAGAGGAGATCTTCATGAGACTTAAG 480  
Db 601 CAGCTGCTCTGGAGTTCACATGATGATGAGAGGAGATCTTCATGAGACTTAAG 660  
Qy 481 TCAAGAAATGTAATTTCTGAAAAATTAATCTTAAATTTGAGATTTTGGAGTTTCTCGA 540  
Db 661 TCAAGAAATGTAATTTCTGAAAAATTAATCTTAAATTTGAGATTTTGGAGTTTCTCGA 720  
Qy 541 CTTCTAATGGGATCTGTGACCTGGCCCAACTTTAACTGGNACTCCCATATATAGT 600

Db 721 CTTCTAATGGGATCTGTGACCTGGCCCAACTTTAACTGGAACTCCCATATATAGT 780  
Qy 601 CTTGAGGCTCTGAAACACCAAGGCTATGACACAAAGTCGGACATCTGTCTCACTGGCATGC 660  
Db 781 CTTGAGGCTCTGAAACACCAAGGCTATGACACAAAGTCGGACATCTGTCTCACTGGCATGC 840  
Qy 661 ATTTTGTATGAGATGCTGTCATGATCATGCAATTCGGTGGCTCCAAATTTCTTATCCAT 720  
Db 841 ATTTTGTATGAGATGCTGTCATGATCATGCAATTCGGTGGCTCCAAATTTCTTATCCAT 900  
Qy 721 GTTTTAAAAAATGTTGAAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAAGAACTA 780  
Db 901 GTTTTAAAAAATGTTGAAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAAGAACTA 960  
Qy 781 AATGCCATCATGGAAGCATGTTGAACCAAGATCTTCAATTAAGACCATCTGCTATCGAA 840  
Db 961 AATGCCATCATGGAAGCATGTTGAACCAAGATCTTCAATTAAGACCATCTGCTATCGAA 1020  
Qy 841 ATTTTAAAAAATCCCTTACCTTGTGAGCAGCTACAGAACCTAAATGTGTAGATATTCAGAA 900  
Db 1021 ATTTTAAAAAATCCCTTACCTTGTGAGCAGCTACAGAACCTAAATGTGTAGATATTCAGAA 1080  
Qy 901 ATGACTCTGGAAGACAAAATTTGGATTTGTGAGAGGAGGCTGCTCATATATTAATGCC 960  
Db 1081 ATGACTCTGGAAGACAAAATTTGGATTTGTGAGAGGAGGCTGCTCATATATTAATGCC 1140  
Qy 961 ATGCAAAAAAGGATCCACCTGCAGACTCTGAGGCGCCTGTGAGAGGATGACAGAAATGACG 1020  
Db 1141 ATGCAAAAAAGGATCCACCTGCAGACTCTGAGGCGCCTGTGAGAGGATGACAGAAATGACG 1200  
Qy 1021 CCAAGAGAAAGGATTCGGCTGAGGAAGCTCCAGGCGGCTGATGAGAAAGCCAGGAAGCTG 1080  
Db 1201 CCAAGAGAAAGGATTCGGCTGAGGAAGCTCCAGGCGGCTGATGAGAAAGCCAGGAAGCTG 1260  
Qy 1081 AAAAGAGTTTGGAGAAAATATGAGAAATAGCAAAATGCAAAATGCAAGATTCAGATCT 1140  
Db 1261 AAAAGAGTTTGGAGAAAATATGAGAAATAGCAAAATGCAAAATGCAAGATTCAGATCT 1320  
Qy 1141 CGGAACCTTTACAGCAGCTGAGTGTGATGTACTCCATGAAAAAACAATTTAAAAAGGAATG 1200  
Db 1321 CGGAACCTTTACAGCAGCTGAGTGTGATGTACTCCATGAAAAAACAATTTAAAAAGGAATG 1380  
Qy 1201 GAAAGAAAGGAGGAGCACTGAGGGAAGACTTTCTGTTTACCCAGGAGGAGGATGAA 1260  
Db 1381 GAAAGAAAGGAGGAGCACTGAGGGAAGACTTTCTGTTTACCCAGGAGGAGGATGAA 1440  
Qy 1261 GAGAGGTGGCAAGGACGAGGAGGAAATCTGATGAACCAACTTTAGAGAACTTCCTGAG 1320  
Db 1441 GAGAGGTGGCAAGGACGAGGAGGAAATCTGATGAACCAACTTTAGAGAACTTCCTGAG 1500  
Qy 1321 TCTCAGCCTATTCCTTCCATGAGCTCCAGAACTTGAATCAATTTGAGAGGATGCCACA 1380  
Db 1501 TCTCAGCCTATTCCTTCCATGAGCTCCAGAACTTGAATCAATTTGAGAGGATGCCACA 1560  
Qy 1381 TCTCAGCCTTGGATACCATG 1399  
Db 1561 TCTCAGCCTTGGATACCATG 1579

## RESULT 8

US-10-751-985-1  
; Sequence 1, Application US/10751985  
; Publication No. US20040126861A1  
; GENERAL INFORMATION:  
; APPLICANT: GUEGLER, Karl et al.  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001007CON  
; CURRENT APPLICATION NUMBER: US/10/751,985  
; CURRENT FILING DATE: 2004-01-07  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: FastSeq for Windows Version 4.0



; SEQ ID NO 1  
; LENGTH: 2869  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-751-985-1

Query Match 96.4%; Score 1397.4; DB 17; Length 2869;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1398; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCTGAAATTCACAGGCGAGCTAAGTGTGTGAGTGTGATCAACAGCCATTTCCACTTAT 60  
DB 181 ATGCTGAAATTCACAGGCGAGCTAAGTGTGTGAGTGTGATCAACAGCCATTTCCACTTAT 240

QY 61 CCAAGACCTTGATTCGAAGAGATACGCTGCTCAACAAAACCTTGGCAGTGGAAATTT 120  
DB 241 CCAAGACCTTGATTCGAAGAGATACGCTGCTCAACAAAACCTTGGCAGTGGAAATTT 300

QY 121 GGAACCTGTCTATCTGCTGTTTCAGACAAAGAGCCAAACGAGGAGAGAAATTAAGGTACTT 180  
DB 301 GGAACCTGTCTATCTGCTGTTTCAGACAAAGAGCCAAACGAGGAGAGAAATTAAGGTACTT 360

QY 181 AAGGAATATCTGTTGGAGAACTAAATCCAAATGAACTGTACAGGCCAAATTTGGAAGCC 240  
DB 361 AAGGAATATCTGTTGGAGAACTAAATCCAAATGAACTGTACAGGCCAAATTTGGAAGCC 420

QY 241 CAATCTCTCCAAAGCTGGACCAACCCAGCCATTTGCAAGTTCCATGCAAGTTTGTGGAG 300  
DB 421 CAATCTCTCCAAAGCTGGACCAACCCAGCCATTTGCAAGTTTCCATGCAAGTTTGTGGAG 480

QY 301 CAAGATAATTTCTGCANTATCAAGGAGTACTGTGAGGCGGAGATCTGGACGATATAAAT 360  
DB 481 CAAGATAATTTCTGCANTATCAAGGAGTACTGTGAGGCGGAGATCTGGACGATATAAAT 540

QY 361 CAGGAATATAAACAAGCTGAAAAATCTTCCAGAAAAATCAAAATATAAGTGGTTATC 420  
DB 541 CAGGAATATAAACAAGCTGAAAAATCTTCCAGAAAAATCAAAATATAAGTGGTTATC 600

QY 421 CAGCTGCTGGAGTTGACTACATGATGAGAGAGAGATATCTCATGAGACTTAAAG 480  
DB 601 CAGCTGCTGGAGTTGACTACATGATGAGAGAGAGATATCTCATGAGACTTAAAG 660

QY 481 TCAAGAGATGATTTCTGAAAAATATCTCTTAAAAATTCGAGATTTGGAGTTTCTCGA 540  
DB 661 TCAAGAGATGATTTCTGAAAAATATCTCTTAAAAATTCGAGATTTGGAGTTTCTCGA 720

QY 541 CTTCTAATGGAGTCTGTGACCTGGCCACAACTTTAACTGGAACTCCCATTTATATGAGT 600  
DB 721 CTTCTAATGGAGTCTGTGACCTGGCCACAACTTTAACTGGAACTCCCATTTATATGAGT 780

QY 601 CTTGAGGCTGTAACACCAAGGCTATGACAAAGTCGACATCTGTCACCTGGCATGC 660  
DB 781 CTTGAGGCTGTAACACCAAGGCTATGACAAAGTCGACATCTGTCACCTGGCATGC 840

QY 661 ATTTTGTATGAGATGTCTGATCAATCATGCAATTCCTGGCTCCAAATTTCTTATCCATT 720  
DB 841 ATTTTGTATGAGATGTCTGATCAATCATGCAATTCCTGGCTCCAAATTTCTTATCCATT 900

QY 721 GTTTTAAAAATTTGTAAGGTGACACACCTTCTCTCTGCTGAGAGATATCAAAAGACTA 780  
DB 901 GTTTTAAAAATTTGTAAGGTGACACACCTTCTCTCTGCTGAGAGATATCAAAAGACTA 960

QY 781 AATGCCATCATGGAAGCATGTTGAAACAGATCTTCAATTAAGACCATCTGCTATCGAA 840  
DB 961 AATGCCATCATGGAAGCATGTTGAAACAGATCTTCAATTAAGACCATCTGCTATCGAA 1020

QY 841 ATTTTAAAAATTCCTTACCTTATGATGAGCAGCTACAGAACTTAATGTGTAGATATCGAA 900  
DB 1021 ATTTTAAAAATTCCTTACCTTATGATGAGCAGCTACAGAACTTAATGTGTAGATATCGAA 1080

QY 901 ATGACTCTGGAGACAAAAATTTGGATTGTGAGAGGAGGCTGCTCATATAAATTAATGCC 960  
DB 1081 ATGACTCTGGAGACAAAAATTTGGATTGTGAGAGGAGGCTGCTCATATAAATTAATGCC 1140

RESULT 9  
US-10-302-172-50  
; Sequence 50, Application US/10302172  
; Publication No. US20040053250A1  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Drmanac, Radje T.  
; TITLE OF INVENTION: No. US20040053250A1 Arginine-rich Protein-like Nucleic Acids and  
; FILE REFERENCE: 803 1CNC  
; CURRENT APPLICATION NUMBER: US/10/302,172  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/225,251  
; PRIOR FILING DATE: 2002-08-20  
; PRIOR APPLICATION NUMBER: PCT US02/05095  
; PRIOR FILING DATE: 2002-03-05  
; PRIOR APPLICATION NUMBER: US 09/799,451  
; PRIOR FILING DATE: 2001-03-05  
; NUMBER OF SEQ ID NOS: 950  
; SOFTWARE: pc\_FL\_genes Version 2.0  
; SEQ ID NO 50  
; LENGTH: 2765  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (597)..(2087)  
US-10-302-172-50

Query Match 95.6%; Score 1385.4; DB 13; Length 2765;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1397; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 ATGCTGAAATTCGAAGGCGAGCTAAGTGTGTGAGTGTGATCAACAGCCATTTCCACTTAT 60  
DB 154 ATGCTGAAATTCGAAGGCGAGCTAAGTGTGTGAGTGTGATCAACAGCCATTTCCACTTAT 213

QY 61 CCAAGACCTTGTATTTCGAAGAGATACGCTGCTTCAACAAAACTTGGCAGTGGAAATTT 120



Db 214 CCAAGACCTTGATTCGAAGAGATACGCTCTTCAACAAAACCTTGGCAGTGGAGTTT 273  
Qy 121 GGAACGTCTATCTGTTTCAGACAAGAACCCAGAGAGAGAAATTAAGGTACTT 180  
Db 274 GGAACGTCTATCTGTTTCAGACAAGAACCCAGAGAGAGAAATTAAGGTACTT 333  
Qy 181 AAGGAATATCTCTGAGAACTTAATCCAAATGAAACTGTACAGCCCAATTTGGAGCC 240  
Db 334 AAGGAATATCTCTGAGAACTTAATCCAAATGAAACTGTACAGCCCAATTTGGAGCC 393  
Qy 241 CAACCTCTCTCAAGCTGGACCCAGCCCAATTTCAAGTTCATCAAGTTTGTGGAG 300  
Db 394 CAACCTCTCTCAAGCTGGACCCAGCCCAATTTCAAGTTCATCAAGTTTGTGGAG 452  
Qy 301 CAGATTAATTTCTGCATTATCAGGAGTACTGTAGGGCCGAGATCTGACGATTAATTT 360  
Db 453 CAGATTAATTTCTGCATTATCAGGAGTACTGTAGGGCCGAGATCTGACGATTAATTT 512  
Qy 361 CAGGAATATAAACAAGCTGGAAAAATCTTTCCAGAAAAATCAATAATAGATGGTTATC 420  
Db 513 CAGGAATATAAACAAGCTGGAAAAATCTTTCCAGAAAAATCAATAATAGATGGTTATC 572  
Qy 421 CAGCTGCTGCTGGAGTGGATCATGCAATGAGAGAGGAGATCTTCATCGAGACTTAAG 480  
Db 573 CAGCTGCTGCTGGAGTGGATCATGCAATGAGAGAGGAGATCTTCATCGAGACTTAAG 632  
Qy 481 TCAAGAATATGATTTCTGAAAAATATCTCTTAAATTTGGAGATTTTGGAGTTTCTCGA 540  
Db 633 TCAAGAATATGATTTCTGAAAAATATCTCTTAAATTTGGAGATTTTGGAGTTTCTCGA 592  
Qy 541 CTTCTAATGGATTCCTGTGACCTGGCCCAACACTTTAACTGGAACTCCCAATATATGAT 600  
Db 693 CTTCTAATGGATTCCTGTGACCTGGCCCAACACTTTAACTGGAACTCCCAATATATGAT 752  
Qy 601 CCGAGGCTCTGAACCAAGCTATGACAAAGTCCGACATCTGTCACCTGGCATCC 560  
Db 753 CCGAGGCTCTGAACCAAGCTATGACAAAGTCCGACATCTGTCACCTGGCATCC 812  
Qy 661 ATTTCGTATGAGATGTGCTGCATGAATCATGATTCGCTGGTCCCAATTTCTATCCAT 720  
Db 813 ATTTCGTATGAGATGTGCTGCATGAATCATGATTCGCTGGTCCCAATTTCTATCCAT 872  
Qy 721 GTTTTAAAAATTTGGAAGTGACACACCTTCTCTCCCTGAGAGATATCCAAAAGAACTA 780  
Db 873 GTTTTAAAAATTTGGAAGTGACACACCTTCTCTCCCTGAGAGATATCCAAAAGAACTA 932  
Qy 781 AATGCCATCATGGAAGAGCTTTGGAACAAGATCTCTTCAATTAAGACCATCTCTATCGAA 840  
Db 933 AATGCCATCATGGAAGAGCTTTGGAACAAGATCTCTTCAATTAAGACCATCTCTATCGAA 992  
Qy 841 ATTTTAAAAATCCCTTACCTTGATGAGCAGCTACAGAACTTAATGTGTAGATATTAGAA 900  
Db 993 ATTTTAAAAATCCCTTACCTTGATGAGCAGCTACAGAACTTAATGTGTAGATATTAGAA 1052  
Qy 901 ATGACTCTGAGACAAAAATTTGGATTTCTCAGAGGAGGCTGCTCATATTAATATGCC 960  
Db 1053 ATGACTCTGAGACAAAAATTTGGATTTCTCAGAGGAGGCTGCTCATATTAATATGCC 1112  
Qy 961 ATGCAAAAAAGATCCACCTGAGACTCTGAGGGCACTCTCAGAGTACAGAAAATGAGC 1020  
Db 1113 ATGCAAAAAAGATCCACCTGAGACTCTGAGGGCACTCTCAGAGTACAGAAAATGAGC 1172  
Qy 1021 CCAAGAGAAAGATGCGGCTGAGAGCTCCAGGGGCTGATGAGAAAGCCAGGAGCTG 1080  
Db 1173 CCAAGAGAAAGATGCGGCTGAGAGCTCCAGGGGCTGATGAGAAAGCCAGGAGCTG 1232  
Qy 1081 AAAAGATTTGGAAGAAAAATATGAAGAAAAATAGCAAAAGCAATGCAAGATTTAGATCT 1140  
Db 1233 AAAAGATTTGGAAGAAAAATATGAAGAAAAATAGCAAAAGCAATGCAAGATTTAGATCT 1292  
Qy 1141 CGGAACTTTTCAGAGCTGAGTGTGATGTATCTCCATGAAAAAACACATTTAAAGGAATG 1200

Db 1293 CGGAATTTTACGACGCTGAGTGTGATGTACTCTCATGAAAAAACACATTTAAAGGAATG 1352  
Qy 1201 GAAGAAAAAGGAGGACCAACCTGAGGAAAGACTTTTCTTGTTCACCCAGGAGAGATGAA 1260  
Db 1353 GAAGAAAAAGGAGGACCAACCTGAGGAAAGACTTTTCTTGTTCACCCAGGAGATGAA 1412  
Qy 1261 GAGAGTGGCAGGAGGAGGAGAAATCTGATGAACCAACTTTTAGAGAACTCGCTGAG 1320  
Db 1413 GAGAGTGGCAGGAGGAGGAGAAATCTGATGAACCAACTTTTAGAGAACTCGCTGAG 1472  
Qy 1321 TCTCAGCCTATTCTCTTCCATGAGCTCCACGAACCTTGAATCAATTTGTAGAGATGCCACA 1380  
Db 1473 TCTCAGCCTATTCTCTTCCATGAGCTCCACGAACCTTGAATCAATTTGTAGAGATGCCACA 1532  
Qy 1381 TCTGACCTTGGATACCATG 1399  
Db 1533 TCTGACCTTGGATACCATG 1551

## RESULT 10

US-10-425-114-16420  
; Sequence 16420, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(533)3/B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 16420  
; LENGTH: 2484  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3063-070-A3\_FLI  
US-10-425-114-16420

Query Match 52.4%; Score 759; DB 13; Length 2484;  
Best Local Similarity 77.5%; Pred. No. 1e-205;  
Matches 1084; Conservative 0; Mismatches 0; Indels 315; Gaps 1;

Qy 1 ATGCTGAAATTCGAAGGAGGAGCTTAAGTGTGTGAGTGGATCAACAGCCATTTCCACTTAT 60  
Db 267 ATGCTGAAATTCGAAGGAGGAGCTTAAGTGTGTGAGTGGATCAACAGCCATTTCCACTTAT 326  
Qy 61 CCAAGACCTTTGATTGCAAGAAGATACGTGCTTCAACAAAAAATTTGGCAGTGGAAAGTTT 120  
Db 327 CCAAGACCTTTGATTGCAAGAAGATACGTGCTTCAACAAAAAATTTGGCAGTGGAAAGTTT 386  
Qy 121 GGAATCTCTATCTGTTTCAGACAAGAACCCAAATGAAATCTGTACAGGCCAATTTGGAAGCC 180  
Db 387 GGAATCTCTATCTGTTTCAGACAAGAACCCAAATGAAATCTGTACAGGCCAATTTGGAAGCC 446  
Qy 181 AAGGAATATCTGTTGAGAACTAAATCCAAATGAAATCTGTACAGGCCAATTTGGAAGCC 240  
Db 447 AAGGAATATCTGTTGAGAACTAAATCCAAATGAAATCTGTACAGGCCAATTTGGAAGCC 506  
Qy 241 CAACCTCTCTCAAGCTGGACCCAGCCCAATTTCAAGTTCCATGCAAGTTTGTGGAG 300  
Db 507 CAACCTCTCTCAAGCTGGACCCAGCCCAATTTCAAGTTCCATGCAAGTTTGTGGAG 566  
Qy 301 CAGATTAATTTCTGCATTATCAGGAGTACTGTGAGGGCCGAGATCTGAGAGATTAATTT 360  
Db 567 CAGATTAATTTCTGCATTATCAGGAGTACTGTGAGGGCCGAGATCTGAGAGATTAATTT 626  
Qy 361 CAGGAATATAAACAAGCTGGAAAAATCTTTCCAGAAAAATCAATAATAGATGGTTTATC 420

627 CAGGAATATAAACAAGCTGGAATAATCTTTCCAGAAAAATCAAAATAAGATGTTTATC 686  
421 CAGCTGCTGCTGGAGTTGACTACATGTCATGAGAGGAGGATCTTCAATGAGACTTAAG 480  
687 CAGCTGCTGCTGGAGTTGACTACATGTCATGAGAGGAGGATCTTCAATGAGACTTAAG 746  
481 TCAAGAATATGTTTCTGAAAAATAATCTCTTAAATTTGGAGATTTTGGAGTTTCTGA 540  
747 TCAAGAATATGTTTCTGAAAAATAATCTCTTAAATTTGGAGATTTTGGAGTTTCTGA 806  
541 CTCTTAATGGATCTGTCGCTGAGCTGAGCAAACTTTAACTGGAATCTCCCATATATAGT 600  
807 CTCTTAATGGATCTGTCGCTGAGCTGAGCAAACTTTAACTGGAATCTCCCATATATAGT 866  
601 CCTGAGGCTCTGAAACACCAAGGCTATGACACAAAGTCGACATCTGGTCACTGGCATGC 660  
867 CCTGAGGCTCTGAAACACCAAGGCTATGACACAAAGTCGACATCTG----- 913  
661 ATTTTGTATGATGCTGCTGATGAATCATGCTATGCTGGCTCCAAATTTCTTATCCATT 720  
914 ----- 913  
721 GTTTTAAATTTGTGAGGTGACACACCTTCTCTCCTGAGAGATATCCAAAGAACTA 780  
914 ----- 913  
781 AATGCCATCATGGAAGCATGTTGAACAAGATCTTTCAATTAAGACCATCTGCTATCGAA 840  
914 ----- 913  
841 ATTTTAAATTCCTTACCTTGTATGAGCAGCTACAGAACCTAATGTGTAGATTCAGAA 900  
914 ----- 913  
901 ATGACTCTGGAAGACAAAAATTTGGATTGTGAGAGGAGGCTGCTCATATTAATATGCC 960  
914 ----- 913  
961 ATGCAAAAAGGATCCACTGCGAGCTCTGAGGGGACTGTGAGAGTACAGAAATGACG 1020  
914 --GCAAAAAGGATCCACTGCGAGCTCTGAGGGGACTGTGAGAGTACAGAAATGACG 971  
1021 CCAAGAAAGGATGCGGCTGAGGAGCTCCAGCGGCTGATCAGAAAGCCAGGAAGCTG 1080  
972 CCAAGAAAGGATGCGGCTGAGGAGCTCCAGCGGCTGATCAGAAAGCCAGGAAGCTG 1031  
1081 AAAAAGATTGTGGAAGAAAAATATGAAGAAAAATAGCAAAAGCAATGCAAGAAATGAGATCT 1140  
1032 AAAAAGATTGTGGAAGAAAAATATGAAGAAAAATAGCAAAAGCAATGCAAGAAATGAGATCT 1091  
1141 CGGAATCTTCAAGAGCTGAGTGTGATGCTATCCATGAAAAACACATTTAAAGGAATG 1200  
1092 CGGAATCTTCAAGAGCTGAGTGTGATGCTATCCATGAAAAACACATTTAAAGGAATG 1151  
1201 GAAGAAAGGAGGAGCAACCTGAGGGAAGACTTTCTTGTTCACCCAGGACGAGGATGAA 1260  
1152 GAAGAAAGGAGGAGCAACCTGAGGGAAGACTTTCTTGTTCACCCAGGACGAGGATGAA 1211  
1261 GAGAGTGGCAAGCAGGGAAGGAACTGTATGAACCACTTTAGAGACCTGCTGAG 1320  
1212 GAGAGTGGCAAGCAGGGAAGGAACTGTATGAACCACTTTAGAGACCTGCTGAG 1271  
1321 TCTCAGCTATTCTTCCATGAGCTCCACGAATCTGAATCAATTTAGAGGATGCCACA 1380  
1272 TCTCAGCTATTCTTCCATGAGCTCCACGAATCTGAATCAATTTAGAGGATGCCACA 1331  
1381 TCTGACCTTGGATACCATG 1399  
1332 TCTGACCTTGGATACCATG 1350

US-09-866-050A-511  
; Sequence 511, Application US/09866050A  
; Publication No. US20030040471A1  
; GENERAL INFORMATION:  
; APPLICANT: Watson, James D.  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Murison, James G.  
; APPLICANT: Kumble, Krishanand D.  
; TITLE OF INVENTION: Compositions Isolated From Skin Cells  
; FILE OF INVENTION: 11000.1011C4U  
; CURRENT APPLICATION NUMBER: US/09/866.050A  
; CURRENT FILING DATE: 2001-05-24  
; NUMBER OF SEQ ID NOS: 725  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 511  
; LENGTH: 2538  
; TYPE: DNA  
; ORGANISM: Rat  
US-09-866-050A-511  
Query Match 46.2%; Score 669.6; DB 10; Length 2538;  
Best Local Similarity 82.2%; Pred. No. 3.8e-180;  
Matches 797; Conservative 0; Mismatches 164; Indels 9; Gaps 2;  
QY 1 ATGTGAAATTCGAAGAGGAGCTAAGTGTGTGAGTGTGATGATCAACAGCCATTTCCACTAT 60  
DB 457 ATGTGAAATTCGAAGAGAGCTGCTAAGT-----GTGACCCACAGCCATTTCTGTGTAC 510  
QY 61 CCAGAGACTTGTATTGCAAGAGATACGTCTTCAACAAAACTTTGGCAGTGGAAAGTTT 120  
DB 511 CACCCACTGTGTATTGCAAGAGATAGTCTTCTCAACAGAGCTTTGACGGGAGTTT 570  
QY 121 GGAAGTGTATCTGTTTTCAGCAAGAAAGCCAAAAGAGAGGAAATTAAGGTTACTT 180  
DB 571 GGAAGTGTATCTGTTTTCAGCAAGAAAGCCAAAGCATGGAGAGAACTAAAAGTACTG 630  
QY 181 AAGGAATATCTGTTGGAGAACTAAATCCAAATCAAACTGTACAGGCCAAATTTGGAAGCC 240  
DB 631 AAGGAATATCTGTTGGAGAAATTAATCCAAATGAAAGTGTGAGGCCAGTATGGAAGCC 690  
QY 241 CAATCTCTCTCAAGCTGGACACCCAGCCATTTCAAGTTTCCATGCAAGTTTGTGAG 300  
DB 691 CAGCTCTCTCCACGCTAAACCAATCCGSCCATCGTCAAGTTTCCACGCAAGCTTCATGGAG 750  
QY 301 CAGGAATATTTCTGATTTATCAGGAGTACTGTGAGGGCCGAGATCTGGAGGATAAAAT 360  
DB 751 CAGGATACATTTTGCATTTATCAGGAGTACTGTGAGGGCCGAGATCTGGACTATAAATC 810  
QY 361 CAGGAATATAAACAAGCTGGAATAATCTTTCCAGAAAAATCAAAATAAGATGTTTATC 420  
DB 811 CAGGAATATAAAGAGCTGGGAAATCTTTCTTCAATCAGATAGTGAATGTTTATC 870  
QY 421 CAGCTGCTGCTGGAGTTGACTATCATGATGAGAGGAGGATCTTCAATGAGACTTAAG 480  
DB 871 CAGTTGCTGCTGGAGTTGATTACATGATGAGAGGAGGATCTTCAATGAGACTTGAA 930  
QY 481 TCAGGAATATGTTTCTGAAAAATAATCTCTTAAATTTGGAGATTTTGGAGTTTCTCGA 540  
DB 931 TCAGGAATATGTTTCTGAAAAATAATCTTACCCAAATTTGGAGTTTGGAGTTTCTCGG 990  
QY 541 CTCTTAATGGATCTGTCGCTGAGCTGAGCAAACTTTAACTGGAATCTCCCATATATAGT 600  
DB 991 CTGCTGATGGGTTTCAATGAGCTGGCTGACAACTCTAAACGGGAGCTCCCATATATAGT 1050  
QY 601 CCGTGGGCTCTGAACACCAAGGCTATGACACAAAGTCGACATCTGCTCACTGGCATGC 660  
DB 1051 CAGAGGCGCTTGAACACCAAGGCTGATGATGCCAAGTCTGACATCTGCTCACTGGCATGC 1110  
QY 661 ATTTTGTATGATGATGCTGCTGATGAATCATGCTATGCTGGCTCCAAATTTCTTATCCATT 720

Db 1111 ATTTTATACGAGATGTGTTGTCATGATCATGGTTTCTTGGCTCCAAATTTCTTGCTCTGTG 1170  
Qy 721 GTTTTAAATTTGTTGAAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTA 780  
Db 1171 GTTTTGAATATGTTGAAGGCAACACGCTTCACTCCCTGACAGATACCCAGAACTA 1230  
Qy 781 AATGCCATCATGGAAGCATGTTGAACAGAAATCCTTTCATTAAGACCATCTGCTATCGAA 840  
Db 1231 AACATCATCATGTACGCGATGTTGAACAGAGTCCCTCCCTGAGACCGCTCGGCTGCAGAG 1290  
Qy 841 ATTTTAAATCCCTTACCTTGATGAGCAGCTACAGAACCTAATGTGTAGATATTCAGAA 900  
Db 1291 ATTTTAAAGCCCTTATGTGAAGAGTGCCTTCAGAGCTGATGTATATAAATCCAGAG 1350  
Qy 901 ATGACTCTCGAAGACAA---AAATTGGATTGTCAGAGGAGCTGCTCATATAATTAAT 957  
Db 1351 GCGACACTGGAAGACNAGAGGAACTCAGCTGTGAGAGGAGGCTGCCCATGCATTAAAC 1410  
Qy 958 GCCATGCAAA 967  
Db 1411 GCCATAAAGA 1420

RESULT 12  
US-10-152-661-511  
; Sequence 511, Application US/10152661  
; Publication No. US2003022835A1  
; GENERAL INFORMATION:  
; APPLICANT: Watson, James D.  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Onrust, Rene  
; APPLICANT: Murison, James G.  
; APPLICANT: Kumble, Krishanand D.  
; TITLE OF INVENTION: Compositions Isolated From Skin Cells  
; FILE REFERENCE: 11000.1011c5  
; CURRENT APPLICATION NUMBER: US/10/152,661  
; PRIOR FILING DATE: 2002-05-20  
; PRIOR APPLICATION NUMBER: 09/866,050  
; PRIOR FILING DATE: 2001-05-24  
; PRIOR APPLICATION NUMBER: 60/221,232  
; PRIOR FILING DATE: 2000-07-25  
; PRIOR APPLICATION NUMBER: 60/206,650  
; PRIOR FILING DATE: 2000-05-24  
; PRIOR APPLICATION NUMBER: 09/312,283  
; PRIOR FILING DATE: 1999-05-14  
; PRIOR APPLICATION NUMBER: PCT/NZ99/00051  
; PRIOR FILING DATE: 1999-04-29  
; PRIOR APPLICATION NUMBER: 09/188,930  
; PRIOR FILING DATE: 1998-11-09  
; PRIOR APPLICATION NUMBER: 09/069,726  
; PRIOR FILING DATE: 1998-04-29  
; NUMBER OF SEQ ID NOS: 725  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 511  
; LENGTH: 2538  
; TYPE: DNA  
; ORGANISM: Rat  
US-10-152-661-511

Query Match 46.2%; Score 669.6; DB 15; Length 2538;  
Best Local Similarity 82.2%; Pred. No. 3.8e-180;  
Matches 797; Conservative 0; Mismatches 164; Indels 9; Gaps 2;  
Qy 1 ATGCTGAAATCCAAAGGAGGAGTAAAGTGTGTGAGTGAATCAACAGCCATTTCCACTTAT 60  
Db 457 ATGCTGAAATCCAAAGAGAGTGTAAAGT-----GTGGACCCACAGCCATTTCTGTGTAC 510  
Qy 61 CCAAGACCTTGATTCAGAGAGATACGCTTCAACAAATTTGGCAGTGGAAAGTTT 120  
Db 511 CCACCCACTGTGATTGCAAGAGATACGTTTCCAAACAGAGCTTGGCAGCGGAGTTT 570

Qy 121 GGAACCTCTATCTGTTTCAGACAAAGCAAGAGAGAGGAAATTTAAAGGTACTT 180  
Db 571 GGAACCTCTATCTGTTTCAGACAAAGCAAGAGAGGAAATTTAAAGGTACTG 630  
Qy 181 AAGAAATATCTGTTGAGAACTAAATCCAAATGAATCTGTACAGGCAATTTGGAAGCC 240  
Db 631 AAGAAATATCTGTTGAGAAATTTAAATCCAAATGAATCTGTACAGGCAATTTGGAAGCC 690  
Qy 241 CAACCTCTCTCCAAAGCTGGACCCAGCCATTTCAAGTTTCCATGCAAGTTTGTGGAG 300  
Db 691 CAGCTCTCTCTCCAAAGCTGGACCCAGCCATTTCAAGTTTCCATGCAAGTTTGTGGAG 750  
Qy 301 CAAGATTAATTTGTCATATACAGGAGTACTGTGAGGCGCAGATCTGGAAGCAATAAAT 360  
Db 751 CAGGTAATTTGTCATATACAGGAGTACTGTGAGGCGCAGATCTGGAAGCAATAAAT 810  
Qy 361 CAGGAATATAAACAAGCTGGAAATCTTTCCAGAAATCAATTAATAGAAATGTTTATC 420  
Db 811 CAGGAATATAAACAAGCTGGAAATCTTTCCAGAAATCAATTAATAGAAATGTTTATC 870  
Qy 421 CAGCTGCTGCTGGAGTTGACTACATGATGAGAGGAGGATCTTTCATGAGACTTTAAG 480  
Db 871 CAGTTGCTGCTGGAGTTGACTACATGATGAGAGGAGGATCTTTCATGAGACTTTGAA 930  
Qy 481 TCAAGAATATGTTCTGAAATTAATCTCTTAAATTTGAGATTTTGAGTTTCTCGA 540  
Db 931 TCAAGAATATGTTCTGAAATTAATCTCTTAAATTTGAGATTTTGAGTTTCTCGG 990  
Qy 541 CTCTAATGGATCTCTGACCTGGCCACAACTTTAACTGGAACCTCCCATTTATATGAT 600  
Db 991 CTGCTGATGGTTTCAATGAGCTGGCTTAACTTAACCCGGACTCCCATTTATATGAT 1050  
Qy 601 CCGAGGCTGTAACACCAAGGCTATGACAAAGTCGAGACATCTGTCACCTGGCATGC 660  
Db 1051 CCAGAGGCTGTAACACCAAGGCTATGACAAAGTCGAGACATCTGTCACCTGGCATGC 1110  
Qy 661 ATTTTGTATGAGATGCTGTCATGAATCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
Db 1111 ATTTTACGAGATGTTGATGAGATGAGTGTGATGAGTGTGATGAGTGTGATGAGTGT 1170  
Qy 721 GTTTTAAATTTGTTGAAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTA 780  
Db 1171 GTTTTGAATATCGTTGAAGGCAACACGCTTCACTCCCTGACAGATACCCAGAACTA 1230  
Qy 781 AATGCCATCATGGAAGCATGTTGAACAGAAATCTTTCATTAAGACCATCTGCTATCGAA 840  
Db 1231 AACATCATCATGTACGCGATGTTGAACAGAGTCCCTCCCTGAGACCGCTCGGCTGCAGAG 1290  
Qy 841 ATTTTAAATPCCCTTACCTTGTATGAGCAGCTACAGAACTTAATGTGTAGATATTCAGAA 900  
Db 1291 ATTTTAAAGCCCTTATGTGGAAGAGTGCCTTCAGCAGCTGATGATATAAATCCAGAG 1350  
Qy 901 ATGACTCTGGAAGACAA---AAATTGGATTGTCAGAGGAGGCTGCTCATATAATTAAT 957  
Db 1351 GCGACACTGGAAGACAAAGAGAACTCAGCTGTGAGAGGAGGCTGCCCATGCATTAAAC 1410  
Qy 958 GCCATGCAAA 967  
Db 1411 GCCATAAAGA 1420

RESULT 13  
US-09-764-875-60  
; Sequence 60, Application US/09764875  
; Publication No. US20040018969A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: P0202  
; CURRENT APPLICATION NUMBER: US/09/764,875  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 1249

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 60
; LENGTH: 451
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (307)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (342)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (347)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (371)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-764-868-241

Query Match      28.7%; Score 415.8; DB 11; Length 451;
Best Local Similarity 98.2%; Pred. No. 5.1e-108;
Matches 439; Conservative 0; Mismatches 6; Indels 2; Gaps 2;

Qy 738 AGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTAAATGCCATCATGGAAG 797
Db 2 AGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTAAATGCCATCATGGAAG 61

Qy 798 CATGTTGAACAAGATCCCTTCATTAAGACCATCTGCTATCGAAATTTTAAAAATCCCTTA 857
Db 62 CATGTTGAACAAGATCCCTTCATTAAGACCATCTGCTATCGAAATTTTAAAAATCCCTTA 121

Qy 858 CTTGTATGACGCTACAGAACCTTAATGTGTAGATATTCAGAAATGACTCTCGAAGACAA 917
Db 122 CTTGTATGACGCTACAGAACCTTAATGTGTAGATATTCAGAAATGACTCTCGAAGACAA 181

Qy 918 AAATTTGGATTGTCAGAGGAGGCTCTCATATTAATGATCCATGCAAAAGGATCCA 977
Db 182 AAATTTGGATTGTCAGAGGAGGCTCTCATATTAATGATCCATGCAAAAGGATCCA 241

Qy 978 CTTGACAGCTCTGAGGGCACTGTTCAGAGTACAGAAATGACCCAGAGAAAGGATCG 1037
Db 242 CTTGACAGCTCTGAGGGCACTGTTCAGAGTACAGAAATGACCCAGAGAAAGGATCG 301

Qy 1038 GCTGAGGAGCTCCAGG-CGGCTGATGAGAAACCCAGGAGCTGAAAGATTTGGAAG 1096
Db 302 GCTGAGGAGCTCCAGGCGGCTGATGAGAAACCCAGGAGCTGAAAGATTTGGAAG 361

Qy 1097 AAAAATATGAAGAAATAGCAACGAAATGCAAGAAATGAGATCTCGAACTTTTCAGCAGC 1156
Db 362 AAAAATGAGAAATAGCAACGAAATGCAAGAAATGAGATCTCGAACTTTTCAGCAGC 421

Qy 1157 TGAGTGTG-ATGTAATCTCCATGAAAA 1182
Db 422 TGAGTGTGTAATGTAATCTCATGAAAAA 448

RESULT 14
US-09-764-868-241
; Sequence 241, Application US/09764868
; Patent No. US2002016871A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 241
; LENGTH: 430
; TYPE: DNA

; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (286)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (286)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (342)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (347)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (371)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-764-868-241

Query Match      27.3%; Score 395.4; DB 9; Length 430;
Best Local Similarity 97.9%; Pred. No. 3.4e-102;
Matches 418; Conservative 1; Mismatches 6; Indels 2; Gaps 2;

Qy 758 CTGAGAGATATCCAAAGAACTAAATGCCATCATGGAAGCAATGTTGAACAAGATCCTT 817
Db 1 CTGAGAGATATCCAAAGAACTAAATGCCATCATGGAAGCAATGTTGAACAAGATCCTT 60

Qy 818 CATTAGACCATCTGCTATCGAAATTTTAAAAATCCCTTACCTTGTGAGCAGCTACAGA 877
Db 61 CATTAGACCATCTGCTATCGAAATTTTAAAAATCCCTTACCTTGTGAGCAGCTACAGA 120

Qy 878 ACCTAAATGTGTAGATATTCAGAAATGACTCTGGAAGCAAAAATTTGGATTGTGAGAAG 937
Db 121 ACCTAAATGTGTAGATATTCAGAAATGACTCTGGAAGCAAAAATTTGGATTGTGAGAAG 180

Qy 938 AGGTGCTCATATTAATTAATGCAATGCAAAAAGGATCCACCTGCGAGACTCTGAGGGCAC 997
Db 181 AGGTGCTCATATTAATTAATGCAATGCAAAAAGGATCCACCTGCGAGACTCTGAGGGCAC 240

Qy 998 TGTGAGAGTACGAAATGACCCGAGAGAAAGGATGCGGCTGAGAGAGCTCCAGG-CG 1056
Db 241 TGTGAGAGTACGAAATGACCCGAGAGAAAGGATGCGGCTGAGAGAGCTCCAGGCGG 300

Qy 1057 GCTGATGAGAAAGCCAGAAAGCTGCAAAAAGATTTGGAAGAAAAATATGAAGAAATAGC 1116
Db 301 GCTGATGAGAAAGCCAGAAAGCTGCAAAAAGATTTGGAAGAAAAATATGAAGAAATAGC 360

Qy 1117 AAAGCAATGCAAGAAATGAGATCTCGAACTTTTCAGCAGCTGAGTGTG-ATGTAATCCA 1175
Db 361 AAAGCAATGCAAGAAATGAGATCTCGAACTTTTCAGCAGCTGAGTGTG-ATGTAATCCA 420

Qy 1176 TGAAGAAA 1182
Db 421 TGAAGAAA 427

RESULT 15
US-09-764-875-376
; Sequence 376, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P202
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 376
; LENGTH: 430
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (286)
```

; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (342)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (347)  
; OTHER INFORMATION: n equals a,t,g, or c  
; NAME/KEY: SITE  
; LOCATION: (371)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-09-764-875-376

Query Match 27.3%; Score 395.4; DB 11; Length 430;  
Best Local Similarity 97.9%; Pred. No. 3.4e-102; Indels 2; Gaps 2;  
Matches 418; Conservative 1; Mismatches 6;  
QY 758 CTGAGAGATATCCAAAGAACTAAATGCCATCATGGAAGCATGTTGAACAGAAATCCTT 817  
Db 1 CTGAGAGATATCCAAAGAACTAAATGCCATCATGGAAGCATGTTGAACAGAAATCCTT 60  
QY 818 CATTAAGACCATCTGCTATCGAAATTTTAAATCCCTTACCTTGATGAGCAGCTACAGA 877  
Db 61 CATTAAGACCATCTGCTATCGAAATTTTAAATCCCTTACCTTGATGAGCAGCTACAGA 120  
QY 878 ACCTAATGTGTAGATATTCAGAAATGACCTCTGGAAGACAAAAATTTGGATTGTCAGAAAG 937  
Db 121 ACCTAATGTGTAGATATTCAGAAATGACCTCTGGAAGACAAAAATTTGGATTGTCAGAAAG 180  
QY 938 AGGCTGCTCATATAATTAATGCGATGCAAAAAAGGATCCACCTGCAGACTCTGAGGGCAC 997  
Db 181 AGGCTGCTCATATAATTAATGCGATGCAAAAAAGGATCCACCTGCAGACTCTGAGGGCAC 240  
QY 998 TGTGAGAAGTACAGAAAAATGACCCAGAGAGAAAGGATGCGCTGAGGAAGCTCCAGG-CG 1056  
Db 241 TGTGAGAAGTACAGAAAAATGACCCAGAGAGAAAGGATGCGCTGANGAAGCTCCAGGGCG 300  
QY 1057 GCTGATGAGAAAGCCAGGAAGCTGAAAAAGATTTGGAAGAAAAATATGAGAAAAATAGC 1116  
Db 301 GCTGATGAGAAAGCCAGGAAGCTGAAAAAGATTTGGAAGAAAAATATGAGAAAAATAGC 360  
QY 1117 AACGGAATCAAGAAATGAGATCTCGAACTTTTCAGCAGCTGAGTGTG-ATGTACTCCA 1175  
Db 361 AACGGAATCAAGAAATGAGATCTCGAACTTTTCAGCAGCTGAGTGTG-ATGTACTCCA 420  
QY 1176 TGAATAA 1182  
Db 421 TGAATAA 427

Search completed: October 3, 2004, 02:52:36  
Job time : 747 secs

---

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2004, 21:17:22 ; Search time 119 Seconds  
(without alignments)  
5757.343 Million cell updates/sec

Title: US-10-803-278-3

Perfect score: 1449

Sequence: 1 atgcctgaattccaagagc.....actggaaaatgaaaaataa 1449

Scoring table: IDENTITY\_NUC

Gapop 10\_0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: \*  
1: /cgn2\_6/prodata/2/ina/5A\_COMB.seq: \*  
2: /cgn2\_6/prodata/2/ina/5B\_COMB.seq: \*  
3: /cgn2\_6/prodata/2/ina/6A\_COMB.seq: \*  
4: /cgn2\_6/prodata/2/ina/6B\_COMB.seq: \*  
5: /cgn2\_6/prodata/2/ina/PCFUS\_COMB.seq: \*  
6: /cgn2\_6/prodata/2/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	146.2	10.1	4263	4	US-09-166-350-29 Sequence 29, Appl
2	143.2	9.9	5448	4	US-09-620-312D-246 Sequence 246, Appl
3	143.2	9.9	5532	4	US-09-620-312D-245 Sequence 245, Appl
4	131.6	9.1	2370	4	US-09-873-404-1 Sequence 1, Appl
5	104.8	7.2	1846	3	US-09-173-581-15 Sequence 15, Appl
6	104.8	7.2	1846	3	US-09-420-915-15 Sequence 15, Appl
7	89	6.1	2119	4	US-09-167-206-1 Sequence 1, Appl
8	88.6	6.1	1060	4	US-09-579-664B-2 Sequence 2, Appl
9	85	5.9	906	3	US-09-221-235-3 Sequence 3, Appl
10	85	5.9	906	3	US-09-221-928-3 Sequence 3, Appl
11	85	5.9	906	3	US-09-221-527-3 Sequence 3, Appl
12	85	5.9	906	3	US-09-221-236-3 Sequence 3, Appl
13	85	5.9	906	3	US-09-221-416-3 Sequence 3, Appl
14	85	5.9	906	3	US-09-221-245-3 Sequence 3, Appl
15	85	5.9	906	3	US-09-163-115-3 Sequence 3, Appl
16	85	5.9	906	3	US-09-221-528-3 Sequence 3, Appl
17	85	5.9	906	3	US-09-593-553-3 Sequence 3, Appl
18	85	5.9	906	3	US-09-221-237-3 Sequence 3, Appl
19	85	5.9	4137	3	US-09-221-235-1 Sequence 1, Appl
20	85	5.9	4137	3	US-09-221-928-1 Sequence 1, Appl
21	85	5.9	4137	3	US-09-221-527-1 Sequence 1, Appl
22	85	5.9	4137	3	US-09-221-236-1 Sequence 1, Appl
23	85	5.9	4137	3	US-09-221-416-1 Sequence 1, Appl
24	85	5.9	4137	3	US-09-221-245-1 Sequence 1, Appl
25	85	5.9	4137	3	US-09-163-115-1 Sequence 1, Appl
26	85	5.9	4137	3	US-09-221-528-1 Sequence 1, Appl
27	85	5.9	4137	3	US-09-593-553-1 Sequence 1, Appl

28	85	5.9	4137	3	US-09-221-237-1	Sequence 1, Appl
29	83	5.7	7218	1	US-08-232-463-14	Sequence 14, Appl
30	82.6	5.7	1251	2	US-09-211-930-2	Sequence 2, Appl
31	82.6	5.7	1251	3	US-09-340-993-2	Sequence 2, Appl
32	82.6	5.7	1251	4	US-09-468-442-2	Sequence 8, Appl
33	82.6	5.7	1353	2	US-09-211-930-8	Sequence 8, Appl
34	82.6	5.7	1353	3	US-09-340-993-8	Sequence 8, Appl
35	82.6	5.7	1353	4	US-09-468-442-8	Sequence 8, Appl
36	82.6	5.7	1542	4	US-09-345-473E-13	Sequence 13, Appl
37	82.6	5.7	3201	2	US-09-211-930-1	Sequence 1, Appl
38	82.6	5.7	3201	3	US-09-340-993-1	Sequence 1, Appl
39	82.6	5.7	3201	4	US-09-468-442-1	Sequence 1, Appl
40	82.6	5.7	3268	4	US-09-688-188B-1	Sequence 1, Appl
41	82.6	5.7	3268	4	US-09-291-417D-1	Sequence 1, Appl
42	80.2	5.5	2121	4	US-09-016-434-1142	Sequence 1142, Ap
43	75.8	5.2	1347	2	US-08-712-709-2	Sequence 2, Appl
44	75.8	5.2	1347	3	US-09-111-444-2	Sequence 2, Appl
45	75.8	5.2	1347	3	US-09-541-228-2	Sequence 2, Appl

#### ALIGNMENTS

##### RESULT 1

US-09-166-350-29

; Sequence 29, Application US/09166350A

; Patent No. 6440663

; GENERAL INFORMATION:

; APPLICANT: Scanlan, Matthew

; APPLICANT: Chen, Yao

; APPLICANT: Stockert, Elisabeth

; APPLICANT: Old, Lloyd

; APPLICANT: Jager, Elke

; APPLICANT: Knuth, Alex

; TITLE OF INVENTION: Renal Cancer Associated Antigens and

; FILE REFERENCE: L0461/7051

; CURRENT APPLICATION NUMBER: US/09/166,350A

; CURRENT FILING DATE: 1998-10-05

; EARLIER APPLICATION NUMBER: US 09/166,350

; EARLIER FILING DATE: 1998-10-05

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: RastSEQ for Windows Version 3.0

; SEQ ID NO 29

; LENGTH: 4263

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-166-350-29

Query Match 10.1%; Score 146.2; DB 4; Length 4263;

Best Local Similarity 52.5%; Pred. No. 3.7e-33;

Matches 372; Conservative 0; Mismatches 328; Indels 9; Gaps 2;

QY	173	AGGTACTTAAGGAATATCTGTTGGAGAACTAATCCAAATGAAACTGTACAGCCCAATT	232
Db	664	ATGTCATCAGGAATTAACATCTCAAGAATGTCGTATAAAGAAAGCAAGAATCAAGGA	723
QY	233	TGGAGCCCAACTCTCTCCAGCTGCACCCAGCCATGTCAAGTTCATGCAAGTT	292
Db	724	GAGAAATGCTGTATGGCAACATGAAGCATCCAAATTTGTCATATAAAGAATCAT	783
QY	293	TTGTGGAGCAAGATAATTTCTGCATTATCAGGAGTACTGTGAGGGCCGAGATCTGAGC	352
Db	784	TTGAAGAAATGGCTCTCTCTACATAGTAAAGAGGAGGATTGTTTGA	843
QY	353	ATAAAATTCAGGAATATAACACAGCTGGAAATCTTTCCAGAAATCAATATAAGAAAT	412
Db	844	AACGAATAAATGCTCAGAAAGGCGCTC-----TGTTTCAAGAGACCAAGATTGGAAT	897
QY	413	GGTTTATCCAGCTGCTGTGGAGTTTGACTACATGATGAGAGGAGATCTTCTCATCGAG	472
Db	898	GGTTTGTGCAGATATGTTTGGCTCTGAGCATGTACATGATAGAAAATTTCTTCCAGG	957

QY 473 ACTTAAGTCAAGAATGATTTCT---GAAAAATATCTCCTTAAATTTGGAGTTTG 529  
 Db 958 ACATAAGTCACAGACATATTTCTAACCAAGATGGGACAGTGCAGCTTGGAGATTTTG 1017  
 QY 530 GAGTTTCTGACATTTCTAATGGGATCCTGTGACCTGGCCCAACCTTTAACTGGAATCTCCOC 589  
 Db 1018 GAATTGCTCGAGTTCTTAATAGTACTGTAGAGCTGCTCGAATTTGCTATAGGCACTCCAT 1077  
 QY 590 ATTATATGAGTCTGAGGCTCTGAACACCAAGGCTATGACACAAAGTCCGACATCTGCT 649  
 Db 1078 ACTACTTGTACCTGAAATCTGTGAAAAACAAGCTTTATAACAATRAAAGTGACATTTGGG 1137  
 QY 650 CACTGCGATGCAATTTGTATGAGATGCTGCGATGAATCATGCTTCGCTGCTCAATT 709  
 Db 1138 CTTTGGGCTGTGCTTCTTATGAGTTGTACATTTAAACATGCAATTTGAAGCTGGAACA 1197  
 QY 710 TCTTATCCATGTTTAAATAATTTGAAGGTGACACACTTCTCTCCCTGAGATATC 769  
 Db 1198 TGAATAACCTGGTACTGAAGATAATCTCCGATCTCTTCTCCAGTGTCTCCACATTACT 1257  
 QY 770 CAAAGCACTAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 829  
 Db 1258 CCTATGATCTCCGAGCTGTCTGCTCAGTTATTTAAAGAAATCTAGGGATAGACCAT 1317  
 QY 830 CTGCTATCGAAATTTTAAATAATCCCTTACCTTGATGAGCAGCTTACAGAA 878  
 Db 1318 CAGTCAACTCCATATTTGGAGAAAGTTTATAGCTAAACGAATCGAAAA 1366

RESULT 2

US-09-620-312D-246  
 ; Sequence 246, Application US/09620312D  
 ; Patent No. 6569662  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tang, Y. Tom  
 ; APPLICANT: Liu, Chenghua  
 ; APPLICANT: Asundi, Vinod  
 ; APPLICANT: Zhang, Jie  
 ; APPLICANT: Ren, Feiyan  
 ; APPLICANT: Chen, Rui-hong  
 ; APPLICANT: Zhao, Qing A.  
 ; APPLICANT: Wehrman, Tom  
 ; APPLICANT: Xue, Aidong J.  
 ; APPLICANT: Yang, Yonghong  
 ; APPLICANT: Wang, Jian-Rui  
 ; APPLICANT: Zhou, Ping  
 ; APPLICANT: Ma, Yunging  
 ; APPLICANT: Wang, Dunrui  
 ; APPLICANT: Wang, Zhiwei  
 ; APPLICANT: John Tillinghast  
 ; APPLICANT: Drmanac, Radoje T.  
 ; TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
 ; TITLE OF INVENTION: Polypeptides  
 ; FILE REFERENCE: 784CIP2B  
 ; CURRENT APPLICATION NUMBER: US/09/620,312D  
 ; PRIOR FILING DATE: 2000-07-19  
 ; PRIOR FILING DATE: 2000-04-25  
 ; PRIOR APPLICATION NUMBER: 09/552,317  
 ; PRIOR FILING DATE: 2000-04-25  
 ; PRIOR APPLICATION NUMBER: 09/488,725  
 ; PRIOR FILING DATE: 2000-01-21  
 ; NUMBER OF SEQ ID NOS: 1105  
 ; SOFTWARE: pt\_FL\_genes Version 1.0  
 ; SEQ ID NO 246  
 ; LENGTH: 5448  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (534)..(4178)  
 ; US-09-620-312D-246

Query Match 9.9%; Score 143.2; DB 4; Length 5448;  
 Best Local Similarity 52.3%; Pred. No. 3.3e-32;

Matches 369; Conservative 0; Mismatches 328; Indels 9; Gaps 2;  
 QY 148 AAAGCCAAACGAGAGAGGAATTAAGGTACTTAAGGAATATCTGTGGAGAACTAAAT 207  
 Db 597 AAATCTCAGAGAATGCGAGACAGTATGTTATCAAGGAATTAACATCTCAAGAATGTCC 656  
 QY 208 CCAATGAACCTGACAGGCGCAATTTGGAAGCCCAACTCTCTCCAGAGTGGACCCCA 267  
 Db 657 AGTAAGAGAGAGAGGAATCAAGGAGAGAGTTCAGCTATGCGAACAATGAGCATCCA 716  
 QY 268 GCCATTGCAAGTCCATGCAAGTTTGTGGAGCAAGATAATTTCTGCATTATCACGGAG 327  
 Db 717 AATATTGTCAGATATAGAGAATCATTTGAAAGAAATGGCTCTCTCTACATAGTAATGAT 776  
 QY 328 TACTGTGAGGCGGAGATCTGGAGGTAATTAAGGTAATTAAGGTAATTAAGGTAATTAAG 387  
 Db 777 TACTGTGAGGCGGAGATCTGGAGGTAATTAAGGTAATTAAGGTAATTAAGGTAATTAAG 830  
 QY 388 TTTCAGAAATCAATTAATAGATGTTTATCCAGCTGCTGCTGGAGTTGACTACATG 447  
 Db 831 TTTCAGAGGATCAGATTTTGGACTGTTGTACAGATATGTTTGGCCCTGAAACATGTA 890  
 QY 448 CATGAGAGGAGTACTTTCATCGAGCTTAAGTCAAGAGATGTAATTTCT---GAAAAAT 504  
 Db 891 CATGATAGAAAAATTTCTTCATCGAGACATTAATTTCTCAGAACATATTTTAACTAAAGAT 950  
 QY 505 AATCTCTTAAATTTGGAGATTTTGGAGATTTTGGAGATTTTGGAGATTTTGGAGATTTT 564  
 Db 951 GGAACAGTACAACTTTGGAGATTTTGGAGATTTTGGAGATTTTGGAGATTTTGGAGAT 1010  
 QY 565 GCCCAACTTTAACTGGAACCTCCCATATATATGAGTCTCGAGGCTGTAACACACAGGC 624  
 Db 1011 GCTGAACTTTGCATAGGAGGCCCATCTACTTGTACCTGAAATCTGTGAAACCAACCT 1070  
 QY 625 TATGACAAAGTCGGACATCTGTGCTACTGGCATGCAATTTTGTATGAGATGCTGTCATG 694  
 Db 1071 TACATAATTAAGTGACATTTGGGCTCTGGGGTGTGCTCTTTATGAGCTGTGTACATT 1130  
 QY 685 AATCATCATTCGCTGCTCCAAATTTCTTATCCATGTTTAAAAATTTGTTGAAGTGAC 744  
 Db 1131 AAACATGCTTTGAAGCTGGCAGTATGAAACCTGCTACTGAAGATAATATCTGGATCT 1190  
 QY 745 ACACCTTCTCTCCCTGAGAGATATCCAAAGAGAACTAAATGCCATCATGGAAGCATGTTG 804  
 Db 1191 TTTCCACTGTGCTTTTGCATTTTCTATGATCTCCGAGTTGGTGTCTCAGTTATTT 1250  
 QY 805 AACAGAAATCTTCATTAAGACCATCTGCTATCGAAATTTTAAAAA 850  
 Db 1251 AAAAGAAATCTAGGGATAGACCATCAGTCACTCCATATTTGGAGA 1296

RESULT 3

US-09-620-312D-245  
 ; Sequence 245, Application US/09620312D  
 ; Patent No. 6569662  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tang, Y. Tom  
 ; APPLICANT: Liu, Chenghua  
 ; APPLICANT: Asundi, Vinod  
 ; APPLICANT: Zhang, Jie  
 ; APPLICANT: Ren, Feiyan  
 ; APPLICANT: Chen, Rui-hong  
 ; APPLICANT: Zhao, Qing A.  
 ; APPLICANT: Wehrman, Tom  
 ; APPLICANT: Xue, Aidong J.  
 ; APPLICANT: Yang, Yonghong  
 ; APPLICANT: Wang, Jian-Rui  
 ; APPLICANT: Zhou, Ping  
 ; APPLICANT: Ma, Yunging  
 ; APPLICANT: Wang, Dunrui  
 ; APPLICANT: Wang, Zhiwei  
 ; APPLICANT: John Tillinghast  
 ; APPLICANT: Drmanac, Radoje T.



; TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
; FILE REFERENCE: Polypeptides  
; CURRENT APPLICATION NUMBER: US/09/620,312D  
; PRIOR FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1105  
; SOFTWARE: pt\_FL\_genes Version 1.0  
; SEQ ID NO 245  
; LENGTH: 5532  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (534)..(4262)  
US-09-620-312D-245

Query Match 9.9%; Score 143.2; DB 4; Length 5532;  
Best Local Similarity 52.3%; Pred. No. 3.4e-32;  
Matches 369; Conservative 0; Mismatches 328; Indels 9; Gaps 2;  
QY 148 AAAGCCAAACGAGGAGGAGGAATTAAGGTACTTAAGGAAATATCTGTTGGAGAACTAAAT 207  
DB 597 AAATCTACAGAAGATGGCAGACAGTATGTTATCAAGGAAATTAACATCTCAAGATGTCC 656  
QY 208 CCAATGAATCTGTACAGGCCAATTTGGAGCCCACTCTCTCAAGCTGGACCAACCA 267  
DB 657 AGTAAAGAAGAAGAAATCAAGGAGAGAAATGTCAGTATTCGCAACATGAAGCATCCA 716  
QY 268 GCCATTGTCAAGTTCATCAAGTTCGAGCAAGATAAATTCGTGATTATCAAGGAG 327  
DB 717 AATATGTCAGTATAGAGAAATCAATTTGAAGAAATGGCTCTCTACATAGTATGAT 776  
QY 328 TACTGTGAGGCGGAGATCTGAGCAGATAAATTCAGGAATATAAACAAGCTGGAAAAATC 387  
DB 777 TACTGTGAGGAGGAGGATCTGTTAAGCGAATAAATGCTCAGAAAGGGGTTT-----TG 830  
QY 388 TTTCAGAAATCAATAATAGATGGTTTATCCAGCTGCTGCTGGAGTTGACTACATG 447  
DB 831 TTTCAGAGATCAGATTTTGGAGCTGTTTACAGATATGTTGGCCCTGAAACATGTA 890  
QY 448 CATGAGGAGGAGTACTTTCATCGAGACTTAAAGTCAAGAAATGTAATTTCT---GAAAAAT 504  
DB 891 CATGATAGAAAAATTTCTTCATCGAGACATTAATCTCAGAACATATTTTAACTAAAGAT 950  
QY 505 AATCTCCTTAAATTTGGAGATTTTGGAGTTTCTGACCTCTAATGGGATCTGTGACCTG 564  
DB 951 GGAACAGTACAACTTGGAGATTTTGGAAATTTGCTAGAGTTCTTAATAGTACTGAGGCTG 1010  
QY 565 GCCACAACTTTAACTGGAACTCCCAATTAATGAGTCTGAGGCTCTGAAACACCAAGGC 624  
DB 1011 GCTCGAATTCATAGGAGCCCACTACTTGTCACTGAAATCTGTGAAACCAACCT 1070  
QY 625 TATGACAAAGTGGAGATCTGGTCACTGGGATGCAATTTTGTATGAGATGCTGGATG 684  
DB 1071 TACAATAATAAAGTGAATTTGGGCTCTGGGCTGTGCTTATGAGCTGTGTACACTT 1130  
QY 685 AATCATGCAATCTGCTGGCTCCAAATTTCTTATCCATTTTAAATTTTAAATTTTGAAGTGC 744  
DB 1131 AATCATGCTTTGAAGCTGGCAGTATGAAGAACCTGCTACTGAGATATATCTGATCT 1190  
QY 745 ACACCTTCTCTCCCTGAGAGATATCAAGAACTAAATGCCATCATGCAAGAGATGTTG 804  
DB 1191 TTTCCACCTGTGTCTTTGCAATTTATCTATGATCTCCGAGTTTGGTGTCTCAGTATTT 1250  
QY 805 AACAGAAATCTTCATTAAGCAATCTGCTAGCAATTTTAAATA 850  
DB 1251 AAAAGAAATCTTAGGATAGACCAATCAGTCAACTCCATATTGGAGA 1296

RESULT 4  
US-09-873-404-1  
; Sequence 1, Application US/09873404  
; Patent No. 6506656  
; GENERAL INFORMATION:  
; APPLICANT: WEBSTER, Marion et al  
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001212-CIP  
; CURRENT APPLICATION NUMBER: US/09/873,404  
; CURRENT FILING DATE: 2001-06-05  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 2370  
; TYPE: DNA  
; ORGANISM: Human  
US-09-873-404-1

Query Match 9.1%; Score 131.6; DB 4; Length 2370;  
Best Local Similarity 51.3%; Pred. No. 5.9e-29;  
Matches 364; Conservative 0; Mismatches 334; Indels 12; Gaps 2;  
QY 175 GTACTTAAGGAAATATCTGTTGGAGAACTAAATCCAAATGAAACTGACAGGCCAATTTG 234  
DB 102 GTCATAAAGAGATCAATTTTGAAGATGCCCATACAGAAAAAGAGCTTCAAGAAA 161  
QY 235 GAAGCCCAACTCTCTCCAGCTGRCACCCAGCCATTCGTCAGTTCATGCAAGTTT 294  
DB 162 GAATGATCTTCTGGAAGAGATGAACATCCCAACATTTAGCCCTTCTCAATTCATTT 221  
QY 295 GTGAGCAAGATATTTCTGCATTATCAGGAGTACTGTGAGGCCGAGATCTGAGCAT 354  
DB 222 CAAGAGATGGAGCTGTTTATGTAATGGAATATTGTGATGGAGGGATCTCATGAAA 281  
QY 355 AAAATTTCAGGAATATAACAGCTGGAAAAATCTTCCAGAAATCAATAATAGAAATGG 414  
DB 282 AGGATCA-----ATAGACAAACCGGGTGTGTATTTAGTGAAGATCAGATCTCGTGGTGG 335  
QY 415 TTTATCAGCTGCTGGGAGTTGACTACATGATCAGAGGAGGATCTTCACTGAGAC 474  
DB 336 TTGTGACAGATCTCTAGGACTAAACATATTATGACAGGAGATATTACACAGGAC 395  
QY 475 TTAAGTCAAGAAATGATTTCTGAAAAATA-----TCTCCTTAAATTTGAGATTTT 528  
DB 396 ATAAAGCTCAGAACATTTTCTTAGCAAGACGGAATGGTGGCAAGCTTGGGAGCTTT 455  
QY 529 GGAGTTTCTGACTCTTAATGGGATCTGTGACTGGCCCACTTTTAACTGGAACCTCC 588  
DB 456 GGTATAGCAGAGTCTGGAATAATTCCATGGAATCTGCTGAACTTGTATTGGAACACT 515  
QY 589 CATTATATGAGTCTGAGGCTCTGAAACACCAAGGCTATGACACAAAGTCCGGACATCTGG 648  
DB 516 TACTACCTGTCCCCAGAGATCTGTCAAGATAAACCCCTACAACTAAACCGGATATTGG 575  
QY 649 TCATGCGATGCATTTTGTATGAGATGTGCTGATGAATCATGCAATTCGCTGGCTCCAAT 708  
DB 576 TCTCTGGCTGTGCTTATATGAGCTCTGCACACTTAAACATCTTTTGGGGTAAACAC 635  
QY 709 TCTTATCCATTTTAAATAATTTGAAGGTGACACACTTCTCTCCTCGAGAGATAT 768  
DB 636 TTACAGCAGCTGGTCTGAGATTTGTCAAGCAGATTTTCCCAATATCTCCGGGGTTT 695  
QY 769 CCAAGAACTAAATGCGATCATGGAAGCATGTTGAACAGAAATCTCTCATTAAGACCA 828  
DB 696 TCTCGTGAAGTCCATTCCTTGATATCTCAGCTCTTTCAGTATCTCTCGAGACCGACCA 755  
QY 829 TCTGCTATCGAAATTTTAAATAATCCCTTACTTCATGAGCAGCTTACAGAA 878  
DB 756 TCCATAAATTTCCATTTTGAAGAGCCCTTTTAGAGATCTTATTCCCAA 805

## RESULT 5

US-09-173-581-15

; Sequence 15, Application US/09173581A

; Patent No. 6013455

; GENERAL INFORMATION:

; APPLICANT: Bandman, Olga

; APPLICANT: Tang, Y. Tom

; APPLICANT: Hillman, Jennifer L.

; APPLICANT: Yue, Henry

; APPLICANT: Guegler, Karl J.

; APPLICANT: Corley, Neil C.

; APPLICANT: Gorgone, Gina

; APPLICANT: Azimzai, Yalda

; APPLICANT: Lu, Aina

; TITLE OF INVENTION: Protein Kinase Homologs

; FILE REFERENCE: PF-0614 US

; CURRENT APPLICATION NUMBER: US/09/173,581A

; CURRENT FILING DATE: 1998-10-15

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PERL Program

; SEQ ID NO 15

; LENGTH: 1846

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE: -

; OTHER INFORMATION: 1567782

US-09-173-581-15

Query Match 7.2%; Score 104.8; DB 3; Length 1846;

Best Local Similarity 53.5%; Pred. No. 5.1e-21;

Matches 242; Conservative 0; Mismatches 207; Indels 3; Gaps 1;

QY 402 AATAATAGATGTTTATCCAGCTGCTGCGAGTTGACTACATGCGAGGAGGAT 461  
DB 369 AATTTGGAGTGGTTGTACAGATATGTTGGCCCTGAAACATGTACATAGAAAAAT 428  
QY 462 ACTTCATCGAGACTTAAAGTCAAAGATGATTTCTGAAAAATAAFTCTCC---TTAAAT 518  
DB 429 TCCTCATCGAGACATTAATCTCAGAACATATTTTAACTAAGATGGAACAGTACAAC 488  
QY 519 TGGAGATTTGGAGTTTCTCGACTTCTAATGGATCTCTGACCTGGCCACAACTTTAAC 578  
DB 489 TGGAGATTTGGAGTTTCTCGACTTCTAATGGATCTCTGACCTGGCCACAACTTTAAC 548  
QY 579 TGGAGATTTGGAGTTTCTCGACTTCTAATGGATCTCTGACCTGGCCACAACTTTAAC 638  
DB 549 AGGACCCCATATCTACTTGTACCTGAAATCTGTGAACAAACCTTACAAATAAAG 608  
QY 639 GGACATCTGGTCTACCTGGCATGCAATTTGTATGAGATGCTGCTGCAATCATGCAATTCGC 698  
DB 609 TGACATTTGGGCTCTGGGCTGTCTCTTTATGAGCTGTGTACACTTAAACATGCTTTGA 668  
QY 699 TGGCTCCAAATTTCTATCCATTTGTTTAAATAATTTGTTGAAGTGACACACCTTCTCTCC 758  
DB 669 AGCTGGCAGTATGAAAACCTGTGATCTGAAGATAATATCTGGATCTTTCCACCTGTCTC 728  
QY 759 TGAGAGATATCCAAAAGAACTAAATCCCATCATGGAAGCATGTTGAAACAAAGATCCTTC 818  
DB 729 TTTGCATTATTCCTATGATCTCCGAGTTTGGTGTCTCAGTTATTTAAAGAAATCCTAG 788  
QY 819 ATTAAGACCATCTGCTATCGAATTTTAAAA 850  
DB 789 GGATAGACCATCATGCTCAACTCCCATATTGGAGA 820

## RESULT 6

US-09-420-915-15

; Sequence 15, Application US/09420915

; Patent No. 6264947

; GENERAL INFORMATION:

; APPLICANT: Bandman, Olga

; APPLICANT: Tang, Y. Tom

; APPLICANT: Hillman, Jennifer L.

; APPLICANT: Yue, Henry

; APPLICANT: Guegler, Karl J.

; APPLICANT: Corley, Neil C.

; APPLICANT: Gorgone, Gina

; APPLICANT: Azimzai, Yalda

; APPLICANT: Lu, Aina

; TITLE OF INVENTION: Protein Kinase Homologs

; FILE REFERENCE: PF-0614 US

; CURRENT APPLICATION NUMBER: US/09/420,915

; CURRENT FILING DATE: 1999-10-20

; EARLIER APPLICATION NUMBER: US 09/173,581

; EARLIER FILING DATE: 1998-10-15

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PERL Program

; SEQ ID NO 15

; LENGTH: 1846

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE: -

; OTHER INFORMATION: 1567782

US-09-420-915-15

Query Match 7.2%; Score 104.8; DB 3; Length 1846;

Best Local Similarity 53.5%; Pred. No. 5.1e-21;

Matches 242; Conservative 0; Mismatches 207; Indels 3; Gaps 1;

QY 402 AATAATAGATGTTTATCCAGCTGCTGCGAGTTGACTACATGCGAGGAGGAT 461  
DB 369 AATTTGGAGTGGTTGTACAGATATGTTGGCCCTGAAACATGTACATAGAAAAAT 428  
QY 462 ACTTCATCGAGACTTAAAGTCAAAGATGATTTCTGAAAAATAAFTCTCC---TTAAAT 518  
DB 429 TCCTCATCGAGACATTAATCTCAGAACATATTTTAACTAAGATGGAACAGTACAAC 488  
QY 519 TGGAGATTTGGAGTTTCTCGACTTCTAATGGATCTCTGACCTGGCCACAACTTTAAC 578  
DB 489 TGGAGATTTGGAGTTTCTCGACTTCTAATGGATCTCTGACCTGGCCACAACTTTAAC 548  
QY 579 TGGAGATTTGGAGTTTCTCGACTTCTAATGGATCTCTGACCTGGCCACAACTTTAAC 638  
DB 549 AGGACCCCATATCTACTTGTACCTGAAATCTGTGAACAAACCTTACAAATAAAG 608  
QY 639 GGACATCTGGTCTACCTGGCATGCAATTTGTATGAGATGCTGCTGCAATCATGCAATTCGC 698  
DB 609 TGACATTTGGGCTCTGGGCTGTCTCTTTATGAGCTGTGTACACTTAAACATGCTTTGA 668  
QY 699 TGGCTCCAAATTTCTATCCATTTGTTTAAATAATTTGTTGAAGTGACACACCTTCTCTCC 758  
DB 669 AGCTGGCAGTATGAAAACCTGTGATCTGAAGATAATATCTGGATCTTTCCACCTGTCTC 728  
QY 759 TGAGAGATATCCAAAAGAACTAAATCCCATCATGGAAGCATGTTGAAACAAAGATCCTTC 818  
DB 729 TTTGCATTATTCCTATGATCTCCGAGTTTGGTGTCTCAGTTATTTAAAGAAATCCTAG 788  
QY 819 ATTAAGACCATCTGCTATCGAATTTTAAAA 850  
DB 789 GGATAGACCATCATGCTCAACTCCCATATTGGAGA 820

## RESULT 7

US-09-167-206-1

; Sequence 1, Application US/09167206A

; Patent No. 6476193

; GENERAL INFORMATION:

; APPLICANT: Nandabalan, Krishnan

; APPLICANT: Schulz, Vincent P.

; APPLICANT: Yang, MeiJa

; TITLE OF INVENTION: NIK1 PROTEIN AND NIK1 PROTEIN COMPLEXES

; FILE REFERENCE: 15966-521 NIK1 protein complexes

; CURRENT APPLICATION NUMBER: US/09/167,206A

; CURRENT FILING DATE: 1998-10-06

; NUMBER OF SEQ ID NOS: 26

; SOFTWARE: Patent In Ver. 2.0

```
; SEQ ID NO 1
; LENGTH: 2119
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (135)..(1469)
US-09-167-206-1

Query Match
Best Local Similarity 6.1%; Score 89; DB 4; Length 2119;
Matches 216; Conservative 0; Mismatches 190; Indels 3; Gaps 1;

QY 461 TACTTCATCGAGACTTAAAGTCAAGAAATGATTTCTGNA---AAATATCTCCTTAAAA 517
D 461 TACTTCATCGAGACTTAAAGTCAAGAAATGATTTCTGNA---AAATATCTCCTTAAAA 517
D 544 TATTGATCGGATCTTAAACAGGCAATGTTTCTGGATGGCAAGCAAAACGCAAGC 603
QY 518 TTGGAGATTTTGGAGTTTCTCCACTTCTAATGGGATCCTGTGACCTGGGCCCAAACTTTAA 577
D 518 TTGGAGATTTTGGAGTTTCTCCACTTCTAATGGGATCCTGTGACCTGGGCCCAAACTTTAA 577
D 604 TTGGAGATTTTGGGCTAGTAGAATTAACCATGACACAGAGTTTTCGCAAAACATTTG 663
QY 578 CTGGAACTCCCATATATAGTCTGAGGCTCTGAACACCAAGGCTATGACAAAGT 637
D 578 CTGGAACTCCCATATATAGTCTGAGGCTCTGAACACCAAGGCTATGACAAAGT 637
D 664 TTGGCAGACCTTATTACATGTCTCCTGAAACAAATGAATCGCATGTCCTACAAATGAGAAAT 723
QY 638 CGGACATCTGGTCACTGGCATCATTTTGTATGAGATGTGTCATGAATCATGCAATTCG 697
D 638 CGGACATCTGGTCACTGGCATCATTTTGTATGAGATGTGTCATGAATCATGCAATTCG 697
D 724 CAGATATCTGGTCACTGGGCTGCTGCTGATGAGTTATGTCATTAATGCTCCTCATTTA 783
QY 698 CTGGCTCCAAATTTCTTATCCATTTGTTTAAATAATTTTGAAGGTGACACACCTTCTCC 757
D 698 CTGGCTCCAAATTTCTTATCCATTTGTTTAAATAATTTTGAAGGTGACACACCTTCTCC 757
D 784 CAGCTTTTAGCCAGAAAGAACTCGCTGGGAAATCAGAGAGGCAAAATTCAGGCGAAATTC 843
QY 758 CTGAGAGATATCCAAAGAACTAAATGCCATCATGGAAGCATGTGAACAGATCTCTT 817
D 758 CTGAGAGATATCCAAAGAACTAAATGCCATCATGGAAGCATGTGAACAGATCTCTT 817
D 844 CATACCGTTACTCTGATGAATTAATGAATTAATTAACGAGGATTTAAACITTAAGGATT 903
QY 818 CATTAAAGCACTCTGTCTGAAATTTTAAATAATCCCTTACCTTGATGA 866
D 818 CATTAAAGCACTCTGTCTGAAATTTTAAATAATCCCTTACCTTGATGA 866
D 904 ACCATCGACCTTCTGTGGAAGAAATTTCTTGAGAACCCITTAATAGAGA 952
QY 904 ACCATCGACCTTCTGTGGAAGAAATTTCTTGAGAACCCITTAATAGAGA 952

RESULT 8
US-09-579-664B-2
; Sequence 2, Application US/09579664B
; Patent No. 6514719
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Bird, Timothy A.
; APPLICANT: Virca, G. Duke
; APPLICANT: Martin, Unja
; APPLICANT: Anderson, Dirk M.
; TITLE OF INVENTION: NOVEL MURINE AND HUMAN KINASES
; FILE REFERENCE: 2923-A
; CURRENT APPLICATION NUMBER: US/09/579,664B
; CURRENT FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1060
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-579-664B-2

Query Match
Best Local Similarity 6.1%; Score 88.6; DB 4; Length 1060;
Matches 269; Conservative 0; Mismatches 279; Indels 3; Gaps 1;

QY 160 GGAGAGAAATTAAGTACTTAAAGAAATATCTGTTGGAGAACTAAATCCAAATGAAGT 219
D 160 GGAGAGAAATTAAGTACTTAAAGAAATATCTGTTGGAGAACTAAATCCAAATGAAGT 219
D 283 GGAGTCCCGTAGCGTTAAAGAAAGTACAGATATTTGATTTAATGGATGCCAAAGCAGT 342
QY 220 GTACAGGCCAATTTGAAGGCCAACTCTCTCCAGCTGGACCCAGCCCAATTTGTCAG 279
```

```
Db 343 GCTGATTGTATCAAGAAATAGACCTCTCTTAAGCAACTCAACCATCCAAATGTAATTA 402
QY 280 TTCCATCAAGATTTTGTGGAGCAAGATAATTTCTGCAATTAACGGAGTACTGTGAGGGC 339
D 403 TACTATGCATCAITTCATTGAGGATAATGAGCTGAACATAGTTTGGAGTTAGCAGATGCT 462
QY 340 CGAGATCTGGACGATAAATTCAGGAATATAACAGCTGGAAATATCTTCCAGAAAT 399
D 463 GGTGACCTCTCCAGAAATGATAAAGCACTTTAAGAAACAAAGAGGCTAAATCCCTGAGAGA 522
QY 400 CAAATAATGAATGCTTTATCCAGCTCTGCTGGAGCTTGAATGAGTCAATGAGAGGAGG 459
D 523 ACCGTTTGGAAATATCTCGTTCAGCTCTGCAATGCACTGACCCACATGCAATTCGAAGA 582
QY 460 ATACTTCATCGAGACTTAAAGTCAAGAAATGATTTCTGAAA---AATAATCTCTTTAA 516
D 583 GTCATGCACAGAGATATAAACCAGCTAATGTGTTCATTACAGCCCATGGGGTAGTAAAA 642
QY 517 ATTGGAGATTTTGGAGTTTCTCGACTTCTAATGGGATCCTGTGACCTGGGCCACAACTTTA 576
D 643 CTGGAGACCTTGGGCTTGGTGGTTCCTGCTCCAAACCAACAGCTGCACATCTTTA 702
QY 577 ACTGGAACTCCCATTAATATGAGTCTCTGAGGCTCTGAAACCAACAGCTATGACAAAG 636
D 703 GTGGGTACACCTTACTACATGTCTCCAGAGAGAAATACATGAAATGATACAACTTCAAG 762
QY 637 TCGAGACATCTGGTCACTGGCATGCAATTTTGTATGAGATGTGCTGCATGATCATGATTC 696
D 763 TCTGACATCTGGTCTCTTGGCTGTCTGATATAGATGGCTGCATGAGAGTCCCTTTC 822
QY 697 GCTGGCTCCAA 707
D 823 TACGGCGACAA 833

RESULT 9
US-09-221-235-3
; Sequence 3, Application US/09221235
; Patent No. 6043040
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,235
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 906
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(906)
US-09-221-235-3

Query Match
Best Local Similarity 5.9%; Score 85; DB 3; Length 906;
Matches 254; Conservative 0; Mismatches 260; Indels 3; Gaps 1;

QY 194 TTGGAGAACTAAATCCAAATGAAACTGTACAGGCCAAATTTGGAGCCCACTCTCTTCCA 253
D 203 TTGATTTAATGGATGCCAAAGACGCTGCTGATTCATCAAGAAATAGATCTTCTTAAAGC 262
QY 254 AGCTGGACCCACCCAGCCATTGTCAAGTTCCATGCAAGTTTGTGGACCAAGATAATTTCT 313
D 263 AACTCAACCATCCAAATGTAATAAATATTAATGATCATCTTCATGAGATAATGAATAA 322
QY 314 GCATTATACCGAGTACTGTGAGGCGCGAGATCTGGACGATGAGCAATAAATTCAGAAATATAAC 373
```

Db 323 ACATAGTTTGGAACTAGCAGATGCTGGCGACCTATCCAGATGATCAAGCATTTTAAAGA 382  
Qy 374 AAGCTGGAAATCTTTCCAGAAATCAAAATATAGAAATGGTTTATCCAGCTGCTGCTGG 433  
Db 383 AGCAAAAGAGGCTAAATCTCTGAAAGAACTGTTTGGAAATATTTTGGTTCAGCTTTCAGTG 442  
Qy 434 GAGTTGACTACATGCGATGAGAGAGAGATCTTCATCGAGACTTAAAGTCAAGAAATGTAT 493  
Db 443 CATTTGAAACATGCAATCTCGAAGAGTCAATGATAGATATAAAACCAAGCTAATGTGT 502  
Qy 494 TTCTGAAA---AATAATCTCTTTAAATTTGGAGATTTTGGAGTTTCTCGACITCTTAATGG 550  
Db 503 TCATTACAGCACTGGGGTGTAAACTTTGGAGATCTTGGCTTGGCCGTTTTCAGCT 562  
Qy 551 GATCCTGTGACCTGGCCACAACTTTAACTGGAATCTCCCATTTATATGAGTCCTGAGGCTC 610  
Db 563 CAAAAACCAAGCTGATGACAAAGTGGACATCTGGTCACTGGCATCATTTTGTATG 670  
Qy 611 TGAACACCAAGCTATGACAAAGTGGACATCTGGTCACTGGCATCATTTTGTATG 670  
Db 623 TACATGAAATGGATCAAACTTCAAACTGACATCTGGTCTCTTGGCTGTCTACTATATG 682  
Qy 671 AGATGCTGCAATGCAATGCAATGCTGCTGGCTCCAA 707  
Db 683 AGATGGCTGCAATGCAAAAGTCTTTCTATGGTGACAA 719  
  
RESULT 10  
US-09-221-928-3  
; Sequence 3, Application US/09221928  
; Patent No. 6121030  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: WNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,928  
; EARLIER FILING DATE: 1998-12-28  
; EARLIER FILING DATE:  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 906  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(906)  
US-09-221-928-3  
  
Query Match 5.9%; Score 85; DB 3; Length 906;  
Best Local Similarity 49.1%; Pred. No. 2.7e-15;  
Matches 254; Conservative 0; Mismatches 260; Indels 3; Gaps 1;  
  
Qy 194 TTGAGAACTAAATCCAAATGAACTGTACAGGCCAATTTGGAAGCCCACTCCTCTCCA 253  
Db 203 TTGATTTAATGGATGCCAAGCAGCTGCTGATTCATCAAGAAATAGATCTTCTTAAGC 262  
Qy 254 AGCTGGACCAACCCAGCCATTTGTCAGAGTTCCATCAAGTTTGGAGCAAGATAATTTCT 313  
Db 263 AACTCAACCACTCCAAATGTAATAAATATTTATGATCATTCATTTGAAGATAAAGAACTAA 322  
Qy 314 GCATTATCACGGAGTACTGTGAGGCGCGAGATCTGACGATAAAATTCAGAAATATAAAC 373  
Db 323 ACATAGTTTGGAACTAGCAGATGCTGGCGACCTATCCAGATGATCAAGCATTTTAAAGA 382  
Qy 374 AAGCTGGAAATCTTTCCAGAAATCAAAATATAGAAATGGTTTATCCAGCTGCTGCTGG 433  
Db 383 AGCAAAAGAGGCTAAATCTCTGAAAGAACTGTTTGGAAATATTTTGTTCAGCTTTGCAATG 442  
Qy 434 GAGTTGACTACATGATGAGAGGAGATCTTCATCGAGACTTTAAAGTCAAAAGATGTAT 493  
Db 443 CATTTGAAACATGCAATCTCGAAGAGTCAATGATAGATATAAAACCAAGCTAATGTGT 502  
Qy 494 TTCTGAAA---AATAATCTCTTTAAATTTGGAGATTTTGGAGTTTCTCGACTTCTAATGG 550  
Db 503 TCATTACAGCACTGGGGTGTAAACTTTGGAGATCTTGGCTTGGCCGTTTTCAGCT 562  
Qy 551 GATCCTGTGACCTGGCCACAACTTTAACTGGAATCTCCCATTTATATGAGTCCTGAGGCTC 610  
Db 563 CAAAAACCAAGCTGCAATCTTTAGTTGGTACGCTTTATTACATGTCTCTCCAGAGAA 622

Qy 494 TTCTGAAA---AATAATCTCTTTAAATTTGGAGATTTTGGAGTTTCTCGACTTCTAATGG 550  
Db 503 TCATTACAGCACTGGGGTGTAAACTTTGGAGATCTTGGCTTGGCCGTTTTCAGCT 562  
Qy 551 GATCCTGTGACCTGGCCACAACTTTAACTGGAATCTCCCATTTATATGAGTCCTGAGGCTC 610  
Db 563 CAAAAACCAAGCTGCAATCTTTAGTTGGTACGCTTTATTACATGTCTCTCCAGAGAA 622  
Qy 611 TGAACACCAAGCTATGACAAAGTGGACATCTGGTCACTGGCATCATTTTGTATG 670  
Db 623 TACATGAAATGGATCAAACTTCAAACTGACATCTGGTCTCTTGGCTGTCTACTATATG 682  
Qy 671 AGATGCTGCAATGCAATGCAATGCTGCTGGCTCCAA 707  
Db 683 AGATGGCTGCAATGCAAAAGTCTTTCTATGGTGACAA 719  
  
RESULT 11  
US-09-221-527-3  
; Sequence 3, Application US/09221527  
; Patent No. 6146832  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: WNI-050  
; CURRENT APPLICATION NUMBER: US/09/221,527  
; EARLIER FILING DATE: 1998-12-28  
; EARLIER FILING DATE:  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 906  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(906)  
US-09-221-527-3  
  
Query Match 5.9%; Score 85; DB 3; Length 906;  
Best Local Similarity 49.1%; Pred. No. 2.7e-15;  
Matches 254; Conservative 0; Mismatches 260; Indels 3; Gaps 1;  
  
Qy 194 TTGAGAACTAAATCCAAATGAACTGTACAGGCCAATTTGGAAGCCCACTCCTCTCCA 253  
Db 203 TTGATTTAATGGATGCCAAGCAGCTGCTGATTCATCAAGAAATAGATCTTCTTAAGC 262  
Qy 254 AGCTGGACCAACCCAGCCATTTGTCAGAGTTCCATCAAGTTTGGAGCAAGATAATTTCT 313  
Db 263 AACTCAACCACTCCAAATGTAATAAATATTTATGATCATTCATTTGAAGATAAAGAACTAA 322  
Qy 314 GCATTATCACGGAGTACTGTGAGGCGCGAGATCTGACGATAAAATTCAGAAATATAAAC 373  
Db 323 ACATAGTTTGGAACTAGCAGATGCTGGCGACCTATCCAGATGATCAAGCATTTTAAAGA 382  
Qy 374 AAGCTGGAAATCTTTCCAGAAATCAAAATATAGAAATGGTTTATCCAGCTGCTGCTGG 433  
Db 383 AGCAAAAGAGGCTAAATCTCTGAAAGAACTGTTTGGAAATATTTTGTTCAGCTTTGCAATG 442  
Qy 434 GAGTTGACTACATGATGAGAGGAGATCTTCATCGAGACTTTAAAGTCAAAAGATGTAT 493  
Db 443 CATTTGAAACATGCAATCTCGAAGAGTCAATGATAGATATAAAACCAAGCTAATGTGT 502  
Qy 494 TTCTGAAA---AATAATCTCTTTAAATTTGGAGATTTTGGAGTTTCTCGACTTCTAATGG 550  
Db 503 TCATTACAGCACTGGGGTGTAAACTTTGGAGATCTTGGCTTGGCCGTTTTCAGCT 562  
Qy 551 GATCCTGTGACCTGGCCACAACTTTAACTGGAATCTCCCATTTATATGAGTCCTGAGGCTC 610  
Db 563 CAAAAACCAAGCTGCAATCTTTAGTTGGTACGCTTTATTACATGTCTCTCCAGAGAA 622

QY 611 TGAACACCAAGCTATGACACAAAGTCGACATCTGGTCACTGGCATGCAATTTGTATG 670  
Db 623 TACATCAAAATGATACAACTTCAATCTGACATCTGCTCTTGGCTGTCTACTATATG 682  
QY 671 AGATGCTGCATGAATCATGATTCGCTGGCTCAA 707  
Db 683 AGATGCTGCATTAACAAAGTCCTTCTATGGTGACAA 719

RESULT 12

US-09-221-236-3  
; Sequence 3, Application US/09221236  
; Patent No. 6146841  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; CURRENT APPLICATION NUMBER: US/09/221,236  
; EARLIER FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 906  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(906)  
US-09-221-236-3

Query Match 5.9%; Score 85; DB 3; Length 906;  
Best Local Similarity 49.1%; Pred. No. 2.7e-15;  
Matches 254; Conservative 0; Mismatches 260; Indels 3; Gaps 1;  
QY 194 TTGGAGAACTAAATCCAAATGAATGTAAGCCCAATTTGGAAGCCCACTCTCTCCA 253  
Db 203 TTGATTAAATGATGCAAGCGTCTGATGCAATCAAGAAATAGATCTCTTAAGC 262  
QY 254 AGCTGGACCAACCCAGCATTGTCAAGTTCATGCAAGTTTGGAGCAAGATAATTTCT 313  
Db 263 AACTCAACCATCCAAATGAATGTAAGCCCAATTTGGAAGCCCACTCTCTCCA 322  
QY 314 GCATTATCAGGAGTACTGTGAGGCGCGAGATCTGGAAGATAATTTGGAAGATAATTTCT 373  
Db 323 ACATAGTTTGGAACTAGCAGATCTCGGACCTATCCAGAAATCATCAAGCATTTTAAGA 382  
QY 374 AAGCTGGAATAATCTTTCCAGAAAATCAATAATAGATGGTTTATCCAGCTGCTGCTGG 433  
Db 383 AGCAAAAGAGGCTAATTTCTGAAAAGAACTGTTGGAAGATATTTTTCAGCTTTGCAGTG 442  
QY 434 GAGTTGACTACATCATGAGAGGAGGATCTTCATCGAGACTTAAAGTCAAGAAATGTAT 493  
Db 443 CATTTGACACATGATCTCGAAGAGTCATGATAGATATATAAACCCAGCTAATGTGT 502  
QY 494 TTCTGAAA---AATAATCTCTTAAATTTGAGATTTTGGAGTTTTCGACTTCTAATGG 550  
Db 503 TCATTACAGCCACTGGGGTGGTAAACTTTGGAGATCTTGGGCTTTGGCCGGTTTTCAGCT 562  
QY 551 GATCCTGTGACCTGCGCACAACTTTAACTGGAATCTCCCATTTATGAGTCTCTGAGGCTC 610  
Db 563 CAAAACCAACAGCTGACATCTTTTAGTTGGTAGCCCTTATTACATGTCTCCAGAGAGAA 622  
QY 611 TGAACACCAAGCTATGACACAAAGTCGGAATCTGGTCACTGGCAGTATTTGTATG 670  
Db 623 TACATGAAAATGGATACAACTTCAATCTGACATCTGGTCTCTTGGCTGTCTACTATATG 682  
QY 671 AGATGCTGCATGAATCATGCAATTCGGCTGGCTCCAA 707  
Db 683 AGATGCTGCATTAACAAAGTCCTTCTATGGTGACAA 719

RESULT 13  
US-09-221-416-3  
; Sequence 3, Application US/09221416  
; Patent No. 6153417  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; CURRENT APPLICATION NUMBER: US/09/221,416  
; EARLIER FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 906  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(906)  
US-09-221-416-3

Query Match 5.9%; Score 85; DB 3; Length 906;  
Best Local Similarity 49.1%; Pred. No. 2.7e-15;  
Matches 254; Conservative 0; Mismatches 260; Indels 3; Gaps 1;  
QY 194 TTGGAGAACTAAATCCAAATGAATGTAAGCCCAATTTGGAAGCCCACTCTCTCCA 253  
Db 203 TTGATTAAATGATGCAAGCGTCTGATGCAATCAAGAAATAGATCTCTTAAGC 262  
QY 254 AGCTGGACCAACCCAGCATTGTCAAGTTCATGCAAGTTTGGAGCAAGATAATTTCT 313  
Db 263 AACTCAACCATCCAAATGAATGTAAGCCCAATTTGGAAGCCCACTCTCTCCA 322  
QY 314 GCATTATCAGGAGTACTGTGAGGCGCGAGATCTGGAAGATAATTTGGAAGATAATTTCT 373  
Db 323 ACATAGTTTGGAACTAGCAGATCTCGGACCTATCCAGAAATCATCAAGCATTTTAAGA 382  
QY 374 AAGCTGGAATAATCTTTCCAGAAAATCAATAATAGATGGTTTATCCAGCTGCTGCTGG 433  
Db 383 AGCAAAAGAGGCTAATTTCTGAAAAGAACTGTTGGAAGATATTTTTCAGCTTTGCAGTG 442  
QY 434 GAGTTGACTACATCATGAGAGGAGGATCTTCATCGAGACTTAAAGTCAAGAAATGTAT 493  
Db 443 CATTTGACACATGATCTCGAAGAGTCATGATAGATATATAAACCCAGCTAATGTGT 502  
QY 494 TTCTGAAA---AATAATCTCTTAAATTTGAGATTTTGGAGTTTTCGACTTCTAATGG 550  
Db 503 TCATTACAGCCACTGGGGTGGTAAACTTTGGAGATCTTGGGCTTTGGCCGGTTTTCAGCT 562  
QY 551 GATCCTGTGACCTGCGCACAACTTTAACTGGAATCTCCCATTTATGAGTCTCTGAGGCTC 610  
Db 563 CAAAACCAACAGCTGACATCTTTTAGTTGGTAGCCCTTATTACATGTCTCCAGAGAGAA 622  
QY 611 TGAACACCAAGCTATGACACAAAGTCGGAATCTGGTCACTGGCAGTATTTGTATG 670  
Db 623 TACATGAAAATGGATACAACTTCAATCTGACATCTGGTCTCTTGGCTGTCTACTATATG 682  
QY 671 AGATGCTGCATGAATCATGCAATTCGGCTGGCTCCAA 707  
Db 683 AGATGCTGCATTAACAAAGTCCTTCTATGGTGACAA 719

RESULT 14

US-09-221-245-3  
; Sequence 3, Application US/09221245  
; Patent No. 6180358  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; EARLIER FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: 09/163,115  
; EARLIER FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 906  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(906)  
US-09-221-245-3

; CURRENT FILING DATE: 1998-12-28  
; EARLIER APPLICATION NUMBER: US 09/163,115  
; EARLIER FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 906  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(906)  
US-09-221-245-3

Query Match 5.9%; Score 85; DB 3; Length 906;

Best Local Similarity 49.1%; Pred. No. 2.7e-15;  
Matches 254; Conservative 0; Mismatches 260; Indels 3; Gaps 1;

```
QY 194 TTGGAGAACTAAATCCAAATGAAACTGTACAGGCCAAATTTGGAAGCCCAACTCCTCTCCA 253
    |||
Db 203 TTGATTTAATGGATGCCAAGACGCTGCTGATTCATCAAGAAATAGATCTCTTAAGC 262

QY 254 AGCTGGACCAACCCAGCCATTGTCAAGTTCATCAAGTTCCTGAGCAAGATAATTTCT 313
    |||
Db 263 AACTCAACCACTCAAAATGTAATAAATAATATGATCATCAATTCATGGAATAATGAATAA 322

QY 314 GCATTATCACGGAGTACTGTGAGGGCCGAGATCTGGACGATATAAATTCAGGAATATAAAC 373
    |||
Db 323 ACATAGTTTTGGAACTAGCAGATGCTGGCGACCTATCCAGAAATGATCAAGCATTTTAAGA 382

QY 374 AAGCTGGAATAATCTTCCAGAAATCAAAATAATAGAAATGGTTTATCCAGCTGCTGTGG 433
    |||
Db 383 AGCAAAAGAGGCTAATTCCTGAAAGAACTGTTTGGAAAGTATTTTGTTCAGCTTTGCAGTG 442

QY 434 GAGTTGACTACATGCTGAGAGGAGATCTTCATCGAGACTTAAAGTCAAAAGATGTAT 493
    |||
Db 443 CATTTGAAACACATGCAATCTCGAAGAGTCAATGATAGAGATATAAAACCGCTAATGTGT 502

QY 494 TTCTGAAA---AATAATCTCTTAAATTTGAGATTTTGGAGTTTCTCGACTTCTAATGG 550
    |||
Db 503 TCATTACAGCCACTGGGGTGTAAACTTGGAGATCTTGGGCTTGGCCGGTTTTCAGCT 562

QY 551 GATCCTGTGACCTGGCCACAACTTAACTGGAACCTCCCATTTATATAGTCTCGAGGCTC 610
    |||
Db 563 CAAAACCAACAGCTGCACATTTCTTAGTTGGTACGCCCTTATTACATGTCTCTCCAGAGAA 622

QY 611 TGAACACCAAGGCTATGACACAAAGTCGAGACATCTGGTCACCTGGCATCATTTTGTATG 670
    |||
Db 623 TACATGAAATGGATACAACTTCAATCTGACATCTGGTCTCTTGGCTGTCTACTATATG 682

QY 671 AGATGTGCTGCATGATCATGATTCGCTGGCTCCAA 707
    |||
Db 683 AGATGGCTGCATTACAAAGTCTTCTATGGTGACAA 719
```

## RESULT 15

US-09-163-115-3  
; Sequence 3, Application US/09163115A  
; Patent No. 6183962  
; GENERAL INFORMATION:  
; APPLICANT: Acton, Susan  
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR  
; FILE REFERENCE: MNI-050  
; CURRENT APPLICATION NUMBER: US/09/163,115A  
; CURRENT FILING DATE: 1998-09-29  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 906  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS

; LOCATION: (1)..(906)  
US-09-163-115-3

Query Match 5.9%; Score 85; DB 3; Length 906;  
Best Local Similarity 49.1%; Pred. No. 2.7e-15;  
Matches 254; Conservative 0; Mismatches 260; Indels 3; Gaps 1;

```
QY 194 TTGGAGAACTAAATCCAAATGAAACTGTACAGGCCAAATTTGGAAGCCCAACTCCTCTCCA 253
    |||
Db 203 TTGATTTAATGGATGCCAAGACGCTGCTGATTCATCAAGAAATAGATCTCTTAAGC 262

QY 254 AGCTGGACCAACCCAGCCATTGTCAAGTTCATCAAGTTCCTGAGCAAGATAATTTCT 313
    |||
Db 263 AACTCAACCACTCAAAATGTAATAAATAATATGATCATCAATTCATGGAATAATGAATAA 322

QY 314 GCATTATCACGGAGTACTGTGAGGGCCGAGATCTGGACGATATAAATTCAGGAATATAAAC 373
    |||
Db 323 ACATAGTTTTGGAACTAGCAGATGCTGGCGACCTATCCAGAAATGATCAAGCATTTTAAGA 382

QY 374 AAGCTGGAATAATCTTCCAGAAATCAAAATAATAGAAATGGTTTATCCAGCTGCTGTGG 433
    |||
Db 383 AGCAAAAGAGGCTAATTCCTGAAAGAACTGTTTGGAAAGTATTTTGTTCAGCTTTGCAGTG 442

QY 434 GAGTTGACTACATGCTGAGAGGAGATCTTCATCGAGACTTAAAGTCAAAAGATGTAT 493
    |||
Db 443 CATTTGAAACACATGCAATCTCGAAGAGTCAATGATAGAGATATAAAACCGCTAATGTGT 502

QY 494 TTCTGAAA---AATAATCTCTTAAATTTGAGATTTTGGAGTTTCTCGACTTCTAATGG 550
    |||
Db 503 TCATTACAGCCACTGGGGTGTAAACTTGGAGATCTTGGGCTTGGCCGGTTTTCAGCT 562

QY 551 GATCCTGTGACCTGGCCACAACTTAACTGGAACCTCCCATTTATATAGTCTCGAGGCTC 610
    |||
Db 563 CAAAACCAACAGCTGCACATTTCTTAGTTGGTACGCCCTTATTACATGTCTCTCCAGAGAA 622

QY 611 TGAACACCAAGGCTATGACACAAAGTCGAGACATCTGGTCACCTGGCATCATTTTGTATG 670
    |||
Db 623 TACATGAAATGGATACAACTTCAATCTGACATCTGGTCTCTTGGCTGTCTACTATATG 682

QY 671 AGATGTGCTGCATGATCATGATTCGCTGGCTCCAA 707
    |||
Db 683 AGATGGCTGCATTACAAAGTCTTCTATGGTGACAA 719
```

Search completed: October 3, 2004, 00:58:23  
Job time : 122 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 2, 2004, 06:03:38 ; Search time 612 Seconds

(without alignments)  
10058.234 Million cell updates/sec

Title: US-10-803-278-3

Perfect score: 1449

Sequence: 1 atgtggaattccaagagc.....actggaaaatgaaaaataa 1449

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002s:\*  
7: Geneseqn2003as:\*  
8: Geneseqn2003bs:\*  
9: Geneseqn2003cs:\*  
10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1399.8	96.6	2080	9	ADC51183 Human cel
2	1399	96.5	1938	4	AAS06746 Polynucle
3	1399	96.5	2552	6	AAD30552 Human kin
4	1399	96.5	2919	9	ADC51185 Human cel
5	1397.4	96.4	2869	6	ABK89295 Human cdn
6	1385.4	95.6	2241	6	ABZ11168 Human pol
7	1179	81.4	2888	9	ADC99143 Human kpp
8	1063	73.4	2631	10	ADE28343 Human kpp
9	759	52.4	2483	9	ADC99142 Human kpp
10	669.6	46.2	2538	6	ABL34984 Rat cdna
11	535	36.9	1071	9	ADC51181 Human cel
12	415.8	28.7	451	4	ABK43470 DNA encod
13	395.4	27.3	430	4	AAS27206 CDNA erco
14	395.4	27.3	430	4	AAK88454 Human dig
15	395.4	27.3	430	4	ABK43786 DNA encod
16	395.4	27.3	430	9	ADB93384 Human cdn
17	336	23.2	336	9	ADC51179 Human cel
18	199	13.7	1438	4	AAI02041 Human rep
19	173.6	12.0	287	6	ABW25194 Human cff
20	172.2	11.9	326014	6	ABK89296 Human gen
21	156	10.8	2022	4	AAI04664 Human rep
22	156	10.8	2022	4	AB197571 Human tes
23	156	10.8	2961	4	AAI04665 Human rep

C

## ALIGNMENTS

RESULT 1  
ADC51183  
ID ADC51183 standard; DNA; 2080 BP.  
XX  
AC ADC51183;  
XX  
DT 18-DEC-2003 (first entry)  
XX  
DE Human cell-cycle related protein coding sequence, SEQ ID 9.  
XX  
KW Human; cytostatic; cell-cycle related protein; nuclear export;  
KW nuclear-cytoplasm transport; cytotoxic; cell-cycle control;  
KW immunological disease; neurological disease; cancer; gene; ds.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 294..1706  
FT /\*tag= a  
FT /product= "Cell-cycle related protein"  
XX  
PN JP2003144168-A.  
XX  
PD 20-MAY-2003.  
XX  
PF 14-NOV-2001; 2001JP-00349158.  
XX  
PR 14-NOV-2001; 2001JP-00349158.  
XX  
PA (KOKU-) KOKURITSU YOBO EISEI KENKYUSHO.  
PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.  
XX  
WPI; 2003-818166/77.  
XX  
P-PSDB; ADC51184.  
XX  
PT Novel DNA or RNA coding a cell-cycle related protein which has nuclear  
PT export function, useful for screening substance that prevent or treat  
PT cell cycle abnormality diseases e.g. immunological disease.  
XX  
PS Disclosure; SEQ ID NO 9; 41pp; Japanese.  
XX  
CC The present invention relates to novel cell-cycle related protein such as  
CC NIMA (Never-In Mitosis, gene A)-related protein kinase of Nek 9 and  
CC coding sequences such as a cell-cycle related protein (ADC51176) having  
CC nuclear export function; cell-cycle related protein (ADC51178) having  
CC nuclear-cytoplasm transport function; cell-cycle related protein

24	156	10.8	2961	4	ABL97572	Ab197572 Human tes
25	149.4	10.3	1781	7	ABX72260	Abx72260 Human NOV
26	149.4	10.3	2257	7	ABX72259	Abx72259 Human NOV
27	148.2	10.2	1588	7	ABX72261	Abx72261 Human NOV
28	147.8	10.2	1510	7	ABZ68734	Abz68734 Nucleotid
29	147.8	10.2	1578	7	ABZ68773	Abz68773 Nucleotid
30	147.8	10.2	1581	7	ABZ68735	Abz68735 Nucleotid
31	147.8	10.2	1638	7	ABZ77150	Abz77150 Human pro
32	147.8	10.2	2050	9	ADC9119	Adc9119 Human kpp
33	147.8	10.2	2059	3	AAC246147	Aaz46147 CDNA sequ
34	147.8	10.2	2102	6	ABZ69800	Abz69800 Human ser
35	147.8	10.2	2110	6	AAZ38850	Aad38850 Human kin
36	147.8	10.2	2220	4	AAK51477	Aak51477 Human pol
37	146.4	10.1	2715	7	ABZ59718	Abz59718 NEK-like
38	146.4	10.1	2715	7	ABZ68774	Abz68774 Nucleotid
39	146.2	10.1	2364	6	ABNS9637	Abns9637 Novel hum
40	146.2	10.1	2372	4	AAK2461	Aak2461 Human pol
41	146.2	10.1	4263	3	AAA09328	Aaa09328 Human can
42	144	9.9	1403	5	AAS90554	Aas90554 DNA encod
43	143.2	9.9	2343	6	ABV75946	Abv75946 Ser/Thr/T
44	143.2	9.9	3645	5	AAS11558	Aas11558 Human cdn
45	143.2	9.9	3729	7	AAI51590	Aai51590 Human ser

(ADC51180) having nuclear export function, cytotoxic function and transfer function in the nucleus; and/or cell-cycle related protein (ADC51182) having transfer function and cytotoxic function in nucleus, nuclear export function and nuclear cytoplasm transport function. The coding sequences for these proteins are given in ADC51175, ADC51177, ADC51179 and ADC51181. The sequences of the invention are useful for screening a substance which promotes or suppresses the transfer function or cytotoxic function in the nucleus, nuclear-cytoplasm transport function and/or a cell-cycle control function. The sequences are also useful for treating or diagnosing an immunological disease, neurological disease or cancer. The present sequence was used to illustrate the invention.

XX SQ Sequence 2080 BP; 652 A; 439 C; 473 G; 516 T; 0 U; 0 Other;

Query Match 96.6%; Score 1399.8; DB 9; Length 2080;  
 Best Local Similarity 99.2%; Pred. No. 0;  
 Matches 1407; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

CC 1 ATGCTGAATTCGAGGAGCTAAGTGTGTGAGTGGATCAACAGCCATTTCCACTTAT 60  
 CC 294 ATGCTGAATTCGAGGAGCTAAGTGTGTGAGTGGATCAACAGCCATTTCCACTTAT 353  
 CC 61 CCAAGACCTTGATTCGAAGAGATACGTGCTTCAACAAAACTTGGCAGTGGAGTTT 120  
 CC 354 CCAAGACCTTGATTCGAAGAGATACGTGCTTCAACAAAACTTGGCAGTGGAGTTT 413  
 CC 121 GGAATCTGTATCTGTTTTCAGCAAGAAAGCAACGAGGAGAGAAATTAAGTACTT 180  
 CC 414 GGAATCTGTATCTGTTTTCAGCAAGAAAGCAACGAGGAGAGAAATTAAGTACTT 473  
 CC 181 ARGGAATCTGTGGAGAACTTAATCCAAATGAACCTGTACGCCCAATTTGGAAGCC 240  
 CC 474 ARGGAATCTGTGGAGAACTTAATCCAAATGAACCTGTACGCCCAATTTGGAAGCC 533  
 CC 241 CCACTCTCTCCAGCTGGACCAACCCAGCCATTGTCAAGTTCCTCAAGTTTTCGGAG 300  
 CC 534 CCACTCTCTCCAGCTGGACCAACCCAGCCATTGTCAAGTTCCTCAAGTTTTCGGAG 593  
 CC 301 CAAGATAATTTCTGCAATATCAGGAGTACTGTGGGCGGAGATCTGGACGATAAAT 360  
 CC 594 CAAGATAATTTCTGCAATATCAGGAGTACTGTGGGCGGAGATCTGGACGATAAAT 653  
 CC 361 CAGGAATATAACAGCTGGAATAATCTTTCCAGAAATCAATTAATAGATGGTTTATC 420  
 CC 654 CAGGAATATAACAGCTGGAATAATCTTTCCAGAAATCAATTAATAGATGGTTTATC 713  
 CC 421 CAGCTCTCTGGAGTGGATGACTACATGATGATGAGAGAGATACCTTCATCGAGACTTAAAG 480  
 CC 714 CAGCTCTCTGGAGTGGATGACTACATGATGATGAGAGAGATACCTTCATCGAGACTTAAAG 773  
 CC 481 TCAGAAAGTATTTCTGGAATAATCTTCTTAAATTTGGAGATTTGGAGTTTCTCGA 540  
 CC 774 TCAGAAAGTATTTCTGGAATAATCTTCTTAAATTTGGAGATTTGGAGTTTCTCGA 833  
 CC 541 CTTCTAATGGATCTCTGACCTGGCCCAACATTTAATGGAATCTCCCATPATATGAT 600  
 CC 834 CTTCTAATGGATCTCTGACCTGGCCCAACATTTAATGGAATCTCCCATPATATGAT 893  
 CC 601 CTTGAGGCTCTGAACACCAAGCTATGACACAAAGTGGACATCTGCTGCGATGC 660  
 CC 894 CTTGAGGCTCTGAACACCAAGCTATGACACAAAGTGGACATCTGCTGCGATGC 953  
 CC 661 ATTTTGTATGAGATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 720  
 CC 954 ATTTTGTATGAGATGCTGATGATGATGATGATGATGATGATGATGATGAT 1013  
 CC 721 GTTTTAAAAATTTGGAAGTGCACACACCTTCTCTCCCTGAGAGATATCCAAAGACTA 780  
 CC 1014 GTTTTAAAAATTTGGAAGTGCACACACCTTCTCTCCCTGAGAGATATCCAAAGACTA 1073  
 CC 781 AATGCCATCATGGAAGCATGTTTGAACCAAGATCTCTTCAATTAAGACCATCTGCTATCGAA 840

Db 1074 AATGCCATCATGGAAGCATGTTTGAACCAAGATCTTTCAATTAAGACCATCTGCTATCGAA 1133  
 Qy 841 ATTTTAAAAATTTCCCTTACCTTGTATGAGCAGCTACAGAACCTTAATGTGTAGATATTCAGAA 900  
 Db 1134 ATTTTAAAAATTTCCCTTACCTTGTATGAGCAGCTACAGAACCTTAATGTGTAGATATTCAGAA 1193  
 Qy 901 ATGACTCTGGAAGACAAAAATTTGGATTTGTGAGAGAGGCTGCTCATATAATTAATGCC 960  
 Db 1194 ATGACTCTGGAAGACAAAAATTTGGATTTGTGAGAGAGGCTGCTCATATAATTAATGCC 1253  
 Qy 961 ATGCAAAAAAGGATCCACCTGCGAGACTCTGAGGCGCTGTGAGAGTACAGAAAAATGACG 1020  
 Db 1254 ATGCAAAAAAGGATCCACCTGCGAGACTCTGAGGCGCTGTGAGAGTACAGAAAAATGACG 1313  
 Qy 1021 CCAAGAGAAAGGATGCGGCTGTGAGAGCTCCAGCGGCTGATCAGAAAAAGCCAGGAAGCTG 1080  
 Db 1314 CCAAGAGAAAGGATGCGGCTGTGAGAGCTCCAGCGGCTGATCAGAAAAAGCCAGGAAGCTG 1373  
 Qy 1081 AAAAGAGTTTGTGAGAGAAATATGAGAAATAGCAAAACGATGCAAGATTTGAGATCT 1140  
 Db 1374 AAAAGAGTTTGTGAGAGAAATATGAGAAATAGCAAAACGATGCAAGATTTGAGATCT 1433  
 Qy 1141 CGGAACCTTTTCAGCAGCTGAGTGTGATGTTACTCCATGAAAAAACACATTTTAAAAAGGAATG 1200  
 Db 1434 CGGAACCTTTTCAGCAGCTGAGTGTGATGTTACTCCATGAAAAAACACATTTTAAAAAGGAATG 1493  
 Qy 1201 GAAGAAAGGAGAGCAACCTGAGGAGAGACTTTTCTGTCACCCAGCAGCAGGATGAA 1260  
 Db 1494 GAAGAAAGGAGAGCAACCTGAGGAGAGACTTTTCTGTCACCCAGCAGCAGGATGAA 1553  
 Qy 1261 GAGAGGTGGCAAGCGGAGAGAGGATCTGATGAACCAACTTTAGAGAACCTGCTGAG 1320  
 Db 1554 GAGAGGTGGCAAGCGGAGAGAGGATCTGATGAACCAACTTTAGAGAACCTGCTGAG 1613  
 Qy 1321 TCTCAGCTTATCTTCCATGAGACCTCCAGAACCTTGAATCAATTTAGAGGATGCCA 1380  
 Db 1614 TCTCAGCTTATCTTCCATGAGACCTCCAGAACCTTGAATCAATTTAGAGGATGCCA 1673  
 Qy 1381 TCTGACCTTCGATACCATGGAGACTGTAACTTAATTTCA 1419  
 Db 1674 TCTGACCTTCGATACCATGGAGACTGTAACTTAATTTCA 1712

RESULT 2  
 AAS06746  
 ID AAS06746 standard; cdna; 1938 BP.  
 XX  
 AC AAS06746;  
 XX  
 DT 12-SEP-2001 (first entry)  
 XX  
 DE Polynucleotide sequence encoding human protein kinase #46.  
 XX  
 KW Human; protein kinase; PK; STK; cancer; cardiovascular disease;  
 KW metabolic disorder; immune related disease; neurological disorder;  
 KW neurodegenerative disorder; inflammatory disorder; infectious disease;  
 KW reproductive disorder; gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FN W0200138503-A2.  
 XX  
 PD 31-MAY-2001.  
 XX  
 PF 22-NOV-2000; 2000WO-US032085.  
 XX  
 PR 24-NOV-1999; 99US-0167482P.  
 XX  
 PA (SUGEN-) SUGEN INC.  
 XX  
 PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;  
 PI Flanagan P, Clary D;  
 XX





lipid disorder; fatty liver; Gaucher's disease; myocardial infarction;  
 drug screening; transgenic animal; antiinflammatory; hepatotropic;  
 hypotensive; anti-HIV; enzyme; ss.  
 Homo sapiens.  
 Key Location/Qualifiers  
 CDS 278..2215  
 /\*tag= a  
 /product= "Human PKIN-5"  
 WO200208399-A2.  
 31-JAN-2002.  
 20-JUL-2001; 2001WO-US023092.  
 21-JUL-2000; 2000US-0220038P.  
 28-JUL-2000; 2000US-0222112P.  
 04-AUG-2000; 2000US-0222831P.  
 11-AUG-2000; 2000US-0224729P.  
 (INCY-) INCYTE GENOMICS INC.  
 (THOR/) THORNTON M.  
 Thornton M, Yue H, Khan FA, Gururajan R, Hafalia AJA, Walia NK,  
 Patterson C, Ramkumar J, Gandhi AR, Policky JL, Baughn MR,  
 Tribouley CM, Bandman O, Nguyen DB, Lu Y, Burford N, Lal P, Ding L,  
 Yao MG, Elliott VS, Recipon SA, Kearney L, Lu DM, Greenwald SR,  
 Tang YT, Xu Y, Walsh RT, Gietzen KU, Yang J, Hillman JL;  
 WPI; 2002-206083/26.  
 P-PSDB; AAEL19147.  
 New human kinase polypeptide, useful in diagnosis, prevention and  
 treatment of cancer, immune disorder, growth and developmental disorder,  
 cardiovascular disorder and lipid disorder.  
 Claim 5; Page 178-179; 196pp; English.  
 The present invention relates to an isolated human kinase polypeptide  
 (PKIN) or a biologically active/immunogenic fragment of PKIN. PKIN is  
 useful for diagnosing, treating and preventing cancer (e.g., leukaemia,  
 lymphoma, melanoma), an immune disorder (e.g., acquired immunodeficiency  
 syndrome (AIDS), Addison's disease, allergy, anaemia, asthma, Crohn's  
 disease, rheumatoid arthritis), a growth and developmental disorder (e.g.  
 bursitis, cirrhosis, hepatitis, psoriasis, Cushing's syndrome), a  
 cardiovascular disorder (e.g., atherosclerosis, hypertension, myocardial  
 infarction), and a lipid disorder (e.g., fatty liver, cholestasis,  
 Gaucher's disease, Niemann-Pick's disease). PKIN is useful in a number of  
 drug screening techniques and to analyse the proteome of a tissue or cell  
 type. PKIN is useful for creating knockin humanised animals or transgenic  
 animals to model human diseases, in somatic or germline gene therapy, to  
 generate a transcript image of a tissue or cell type, for detecting  
 differences in the chromosomal location due to translocation, inversion,  
 etc., among normal, carrier or affected individuals, and as hybridisation  
 probes for mapping naturally occurring genomic sequences. PKIN is useful  
 in southern or northern analysis, dot blot or other membrane-based  
 technologies, in PCR technologies, in dipstick, pin, multiformat enzyme  
 linked immunosorbent (ELISA)-like assays and in microarrays utilising  
 fluids or tissues from patients to detect altered PKIN expression. The  
 present sequence is human PKIN-5 cDNA  
 Sequence 2552 BP; 813 A; 516 C; 608 G; 615 T; 0 U; 0 Other;  
 Query Match 96.5%; Score 1399; DB 6; Length 2552;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 ATGCTGAATTCACAGAGCAGCTAAGTGTGTGAGTGGATCAACAGCCATTCCACTTAT 60  
 278 ATGCTGAATTCACAGAGCAGCTAAGTGTGTGAGTGGATCAACAGCCATTCCACTTAT 337

QY 51 CCAGAGCCTTGATTGTCGAGGAGATAGTCTCTCAACAAAAAAGCTTGGAGTGGAGTTT 120  
 DB 338 CCAAGAGCCTTGATTGTCGAGGAGATAGTCTCTCAACAAAAAAGCTTGGAGTGGAGTTT 397  
 QY 121 GGAACTGTCTATCTGTTTTCAGACAAAGAAAGCCAAACGAGGAGAGGAATTAAGAGTACTT 180  
 DB 398 GGAACTGTCTATCTGTTTTCAGACAAAGAAAGCCAAACGAGGAGAGGAATTAAGAGTACTT 457  
 QY 181 AAGGAATATCTGTTGGAGAACTAAATCCAAATGAATGTACAGGCCAATTTGGAGGCC 240  
 DB 458 AAGGAATATCTGTTGGAGAACTAAATCCAAATGAATGTACAGGCCAATTTGGAGGCC 517  
 QY 241 CAATCTCTCTCAAGCTGGACCCAGCCATCTCAAGTTCCATGCAAGTTTGTGGAG 300  
 DB 518 CAATCTCTCTCAAGCTGGACCCAGCCATCTCAAGTTCCATGCAAGTTTGTGGAG 577  
 QY 301 CAAGATAATTTCTGCAATATCACGGAGTATCTGTGAGGGCCGAGATCTGGACGATAAAAT 360  
 DB 578 CAAGATAATTTCTGCAATATCACGGAGTATCTGTGAGGGCCGAGATCTGGACGATAAAAT 637  
 QY 361 CAGGAATATATAAACAAGCTGGAAATAATCTTCCAGAAATCAATAATAGAAATGTTTATC 420  
 DB 638 CAGGAATATATAAACAAGCTGGAAATAATCTTCCAGAAATCAATAATAGAAATGTTTATC 697  
 QY 421 CAGCTGCTGCTGGAGTGTACTACATGATGAGGAGGAGATCTTCAATCGAGACTTAAAG 480  
 DB 698 CAGCTGCTGCTGGAGTGTACTACATGATGAGGAGGAGATCTTCAATCGAGACTTAAAG 757  
 QY 481 TCRAAGATGCTATTTCTGAAATAATCTCTTAATAATTTGGAGATTTTGGAGTTTCTCGA 540  
 DB 758 TCRAAGATGCTATTTCTGAAATAATCTCTCTTAATAATTTGGAGATTTTGGAGTTTCTCGA 817  
 QY 541 CTTCTAATGGGATCCTCTGACCTGGCCACAACTTTAACTCGAACTTCCCAATTTATGAGT 600  
 DB 818 CTTCTAATGGGATCCTCTGACCTGGCCACAACTTTAACTCGAACTTCCCAATTTATGAGT 877  
 QY 601 CCTGAGGCTCTGAACACCAAGCTATGACACAAAGTCGACATCTGGTCACTGGGATGC 660  
 DB 878 CCTGAGGCTCTGAACACCAAGCTATGACACAAAGTCGACATCTGGTCACTGGGATGC 937  
 QY 661 ATTTTGTATGAGATGCTCTGATGATCATGATCATGCTGCTGCTCCCAATTTCTTATCCATT 720  
 DB 938 ATTTTGTATGAGATGCTCTGATGATCATGATCATGCTGCTGCTCCCAATTTCTTATCCATT 997  
 QY 721 GTTTTAAAAATTTGTTGAAGTGTGACACACTTCTCTCCCTGAGAGATATCCAAAAGAACTA 780  
 DB 998 GTTTTAAAAATTTGTTGAAGTGTGACACACTTCTCTCCCTGAGAGATATCCAAAAGAACTA 1057  
 QY 781 AATGCCATCATGAAAGAGCTGTTGAAACAAAGATCCTTTCATTAGACCATCTGCTATCGAA 840  
 DB 1058 AATGCCATCATGAAAGAGCTGTTGAAACAAAGATCCTTTCATTAGACCATCTGCTATCGAA 1117  
 QY 841 ATTTTAAAAATTCCTTACCTTGTATGAGCAGCTACAGAACTTAATGTAGATATTCAGAA 900  
 DB 1118 ATTTTAAAAATTCCTTACCTTGTATGAGCAGCTACAGAACTTAATGTAGATATTCAGAA 1177  
 QY 901 ATGACTCTGGAACACAAAAATTTGGATTGTGACAGAGGCTGCTCATATAATATATGCC 960  
 DB 1178 ATGACTCTGGAACACAAAAATTTGGATTGTGACAGAGGCTGCTCATATAATATATGCC 1237  
 QY 961 ATGCAAAAAAGGATCCACCTGACAGCTCTGAGGGCCTCTGTGAGGAGCTGTGAGGAGCTG 1020  
 DB 1238 ATGCAAAAAAGGATCCACCTGACAGCTCTGAGGGCCTGTGAGGAGCTGTGAGGAGCTG 1297  
 QY 1021 CCAAGAGAAAGGATGGGCTGAGGAGAGCTCCAGGGCTGTGAGGAGAGCCAGGAGCTG 1080  
 DB 1298 CCAAGAGAAAGGATGGGCTGAGGAGAGCTCCAGGGCTGTGAGGAGAGCCAGGAGCTG 1357  
 QY 1081 AAAAAGATTTCTGGAAGAAAAATATGAAGAAAAATAGCAAAACGAATGCAAGAAATTTGAGATCT 1140  
 DB 1358 AAAAAGATTTCTGGAAGAAAAATATGAAGAAAAATAGCAAAACGAATGCAAGAAATTTGAGATCT 1417  
 QY 1141 CGGAACCTTTGAGGAGCTGAGTGTGATGTATCTCCATGAAAAAACACATTTAAAGGAATG 1200

Db 1418 CGGAACCTTTCCAGCAGCTGAGTGTGATCTCTCCATGAAACCAACATTTAAAGGAATG 1477  
 QY 1201 GAAGAAAGAGGAGCAACCTGAGGAGCACTTCTGTGTCACCCAGGAGGAGTGA 1260  
 Db 1478 GAAGAAAGAGGAGCAACCTGAGGAGCACTTCTGTGTCACCCAGGAGGAGTGA 1537  
 QY 1261 GAGAGTGGCAAGGAGGAGCAACCTGATGAAACCACTTTAGAGAACCTGCTCGAG 1320  
 Db 1538 GAGAGTGGCAAGGAGGAGCAACCTGATGAAACCACTTTAGAGAACCTGCTCGAG 1597  
 QY 1321 TCTCAGCCTATTCTCTCCATGAGCACTCCAGCACTTGAATCAATTTGAGAGATGCCACA 1380  
 Db 1598 TCTCAGCCTATTCTCTCCATGAGCACTCCAGCACTTGAATCAATTTGAGAGATGCCACA 1657  
 QY 1381 TCTGACCTTTGGATACCATTG 1399  
 Db 1658 TCTGACCTTTGGATACCATTG 1676

RESULT 4

AD511185  
 ID AD511185 standard; DNA; 2939 BP.  
 AC AD511185;  
 XX  
 DT 18-DEC-2003 (first entry)  
 DE Human cell-cycle related protein coding sequence, SEQ ID 11.  
 XX  
 DE Human; cytostatic; cell-cycle related protein; nuclear export;  
 KW nuclear-cytoplasm transport; cytotoxic; cell-cycle control;  
 KW immunological disease; neurological disease; cancer; gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 294..2231  
 FT /\*tag= a  
 FT /product= "Cell-cycle related protein"  
 XX  
 JP2003144168-A.  
 PD 20-MAY-2003.  
 XX  
 PF 14-NOV-2001; 2001JP-00349158.  
 XX  
 PR 14-NOV-2001; 2001JP-00349158.  
 XX  
 PA (KOKU-) KOKURITSU YOBO EISEI KENKYUSHO.  
 PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.  
 XX  
 WIPI; 2003-818166/77.  
 P-PSDB; AD511186.  
 XX

Novel DNA or RNA coding a cell-cycle related protein which has nuclear export function, useful for screening substance that prevent or treat cell cycle abnormality diseases e.g. immunological disease.  
 PS Disclosure; SEQ ID NO 11; 41pp; Japanese.  
 XX  
 CC The present invention relates to novel cell-cycle-related protein such as NIMA (Never-In Mitosis, gene A)-related protein kinase of Nek 9 and coding sequences such as a cell-cycle related protein (AD511176) having nuclear export function; cell-cycle related protein (AD511178) having nuclear-cytoplasm transport function; cell-cycle related protein and (AD511180) having nuclear export function, cytotoxic function and transfer function in the nucleus; and/or cell-cycle related protein (AD511182) having transfer function and cytotoxic function in nucleus, nuclear export function and nuclear cytoplasm transport function. The coding sequences for these proteins are given in AD511175, AD511177, AD511179 and AD511181. The sequences of the invention are useful for screening a substance which promotes or suppresses the transfer function

or cytotoxic function in the nucleus, nuclear-cytoplasm transport function and/or a cell-cycle control function. The sequences are also useful for treating or diagnosing an immunological disease, neurological disease or cancer. The present sequence was used to illustrate the invention.  
 XX  
 SQ Sequence 2939 BP; 944 A; 600 C; 675 G; 720 T; 0 U; 0 Other;  
 Query Match 96.5%; Score 1399; DB 9; Length 2939;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATGCTGAAATCCAAAGAGGAGCTAAGTGTGTGAGTGCATCAACAGCCATTTCCACTAT 60  
 Db 294 ATGCTGAAATCCAAAGAGGAGCTAAGTGTGTGAGTGCATCAACAGCCATTTCCACTAT 353  
 QY 61 CCAAAGACCTTGATTGCAAGAGATACGTGCTTCAACAAAACCTTGGCAGTGAAGTTT 120  
 Db 354 CCAAAGACCTTGATTGCAAGAGATACGTGCTTCAACAAAACCTTGGCAGTGAAGTTT 413  
 QY 121 GGAACCTGTATCTGTGTTTCAACAAGAACCCAAAGAGGAGGAGGATTAAGGTACTT 180  
 Db 414 GGAACCTGTATCTGTGTTTCAACAAGAACCCAAAGAGGAGGAGGATTAAGGTACTT 473  
 QY 181 AAGGAATATCTGTGGAGAACTAAATCCAAATGAACTGTACAGGCCAATTTGGAAGCC 240  
 Db 474 AAGGAATATCTGTGGAGAACTAAATCCAAATGAACTGTACAGGCCAATTTGGAAGCC 533  
 QY 241 CAACCTCTTCCAAGCTGGACCCAGCCATTCCTCAAGTTCCATGCAAGTTTGTGGAG 300  
 Db 534 CAACCTCTTCCAAGCTGGACCCAGCCATTCCTCAAGTTCCATGCAAGTTTGTGGAG 593  
 QY 301 CAAGATAATTTCTGCATTTACCGAGTACTGTGAGGCGGAGATCTGGAGATTAAGT 360  
 Db 594 CAAGATAATTTCTGCATTTACCGAGTACTGTGAGGCGGAGATCTGGAGATTAAGT 653  
 QY 361 CAGGAATATAAACAGCTGGAAAAATCTTCCAGAAAAATCAAAATATAGAAATGTTTATC 420  
 Db 654 CAGGAATATAAACAGCTGGAAAAATCTTCCAGAAAAATCAAAATATAGAAATGTTTATC 713  
 QY 421 CAGCTGCTGTGGAGTTGACTACATGATGAGAGGAGGATCTTCATCGAGACTTAAG 480  
 Db 714 CAGCTGCTGTGGAGTTGACTACATGATGAGAGGAGGATCTTCATCGAGACTTAAG 773  
 QY 481 TCAAGAAATGTATTTCTGAAAAATTAATCTCCTTAAATTTGAGATTTTCTCGA 540  
 Db 774 TCAAGAAATGTATTTCTGAAAAATTAATCTCCTTAAATTTGAGATTTTCTCGA 833  
 QY 541 CTTCTAATGGATCCTGTGACCTGGCCACAACCTTTAACTGGAACTCCCATTTATGAT 600  
 Db 834 CTTCTAATGGATCCTGTGACCTGGCCACAACCTTTAACTGGAACTCCCATTTATGAT 893  
 QY 601 CTTGAGGCTCTGAAACCAAGGCTATGACAAAGTGGACATCTGGTCACTGGCATGC 660  
 Db 894 CTTGAGGCTCTGAAACCAAGGCTATGACAAAGTGGACATCTGGTCACTGGCATGC 953  
 QY 661 ATTTTGTATGAGATGTGTCATGAATCATGCTCGCTCCCAATTTCTTATCCATT 720  
 Db 954 ATTTTGTATGAGATGTGTCATGAATCATGCTCGCTCCCAATTTCTTATCCATT 1013  
 QY 721 GTTTTAAAAATTTGAAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTA 780  
 Db 1014 GTTTTAAAAATTTGAAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTA 1073  
 QY 781 AATGCCATCATGGAAGCATGTGAAACAAGAAATCTTCAATTAAGACCATCTGCTATCGAA 840  
 Db 1074 AATGCCATCATGGAAGCATGTGAAACAAGAAATCTTCAATTAAGACCATCTGCTATCGAA 1133  
 QY 841 ATTTTAAAAATCCCTTACTTGTAGAGCAGCTACAGAACCTTAATGTGTAGATATTCAGAA 900  
 Db 1134 ATTTTAAAAATCCCTTACTTGTAGAGCAGCTACAGAACCTTAATGTGTAGATATTCAGAA 1193  
 QY 901 ATGACTCTTGGAGACAAAATTTGGATTGTGAGAGGAGGCTGCTCATATAATATGCC 960

Db 1194 ATGACTCTGGAGACAAAATTTGGATTGTCAGAGGAGGCTGCTATATTAATGCC 1253  
Qy 961 ATGCAAAAAGGATCCACCTGACAGCTCTCAGGCACTGTGCAAGTACAGAAAATGAGC 1020  
Db 1254 ATGCAAAAAGGATCCACCTGACAGCTCTCAGGCACTGTGCAAGTACAGAAAATGAGC 1313  
Qy 1021 CCAAGAGAAAGGATGCGCTCAGGAAGCTCCAGCGCTGATGAGAAAGCCAGGAGCTG 1080  
Db 1314 CCAAGAGAAAGGATGCGCTCAGGAAGCTCCAGCGCTGATGAGAAAGCCAGGAGCTG 1373  
Qy 1081 AAAAAGATTGTGGAAGAAAATATGAAGAAAATAGCAAAATGCAAAATGAGATCT 1140  
Db 1374 AAAAAGATTGTGGAAGAAAATATGAAGAAAATAGCAAAATGCAAAATGAGATCT 1433  
Qy 1141 CGGAACCTTCAGCACTGAGTGTGATGTTACTCCATGAAAAACACATTTAAAGGAATG 1200  
Db 1434 CGGAACCTTCAGCACTGAGTGTGATGTTACTCCATGAAAAACACATTTAAAGGAATG 1493  
Qy 1201 GAAGAAAAGGAGGACCACTGAGGGAAGACTTCTGTTCCACCCAGGACGAGGATGAA 1260  
Db 1494 GAAGAAAAGGAGGACCACTGAGGGAAGACTTCTGTTCCACCCAGGACGAGGATGAA 1553  
Qy 1261 GAGAGTGGCAAGCAGGGAAGGAAATCTGTATGAACCACTTTAGAGAACTTGCCTGAG 1320  
Db 1554 GAGAGTGGCAAGCAGGGAAGGAAATCTGTATGAACCACTTTAGAGAACTTGCCTGAG 1613  
Qy 1321 TCTCAGCTATTCTCTCCATGGACCTCCAGCACTTGAATCAATTTAGAGGATGCCACA 1380  
Db 1614 TCTCAGCTATTCTCTCCATGGACCTCCAGCACTTGAATCAATTTAGAGGATGCCACA 1673  
Qy 1381 TCTGACCTTGGATACCATG 1399  
Db 1674 TCTGACCTTGGATACCATG 1692

## RESULT 5

ABK89295  
ID ABK89295 standard; cDNA; 2869 BP.

AC ABK89295;

DT 21-OCT-2002 (first entry)

XX Human cDNA encoding novel serine/threonine serine kinase.

DE Human; ss; gene; serine/threonine protein kinase; inflammation; cancer;  
KW arteriosclerosis; psoriasis; SNF kinase; transgenic; chromosome 3.

XX Homo sapiens.

XX Key Location/Qualifiers  
FH CDS 181..2118  
FT /\*tag= a  
FT /product= "Serine/threonine protein kinase"

XX US2002082189-A1.

XX 27-JUN-2002.

XX 07-DEC-2000; 2000US-00731231.

XX 07-DEC-2000; 2000US-00731231.

XX (GUEG/) GUEGLER K.

XX (KETCH/) KETCHUM K. A.

XX (DFRA/) DI FRANCESCO V.

XX (BEAS/) BEASLEY E M.

XX Guegler K, Ketchum KA, Di Francesco V, Beasley EM;

XX WPI; 2002-598989/64.

XX P-PSDB; ABG31081.

XX New isolated human kinase peptide for detecting a modulator of the  
PT peptide's expression, activity or function, that can be used to treat  
XX disorders or disease.  
XX Claim 4; Fig 1; 321pp; English.  
XX The invention relates to an isolated kinase peptide, comprising, a novel  
CC human serine/threonine protein kinase sequence, its allelic variant or  
CC orthologue, where the variant is encoded by a nucleic acid molecule that  
CC hybridises under stringent conditions to the opposite strand of kinase  
CC cDNA or gene, or a fragment comprising 10 contiguous amino acids. Also  
CC included are an antibody that selectively binds the kinase, a gene chip  
CC comprising a nucleic acid (or its complement) which encodes the kinase, a  
CC transgenic non-human animal comprising the nucleic acid, a nucleic acid  
CC vector comprising the nucleic acid and a host cell comprising the vector.  
CC The kinase is used to identify a modulator of the kinase or to identify  
CC an agent that binds to the kinase, which can be used to treat a disease  
CC or condition e.g. inflammation, cancer, arteriosclerosis and psoriasis.  
CC The nucleic acid encoding the kinase is used to produce the kinase. A  
CC detection agent is used to detect the kinase and an oligonucleotide is  
CC used to detect a nucleic acid encoding the kinase in a sample. The kinase  
CC can be used: (a) to raise antibodies against the kinase or to elicit  
CC another immune response; (b) as a reagent (including a labeling reagent)  
CC in assays to determine levels of a kinase (or its binding partner or  
CC ligand) in biological fluids; and (c) as markers for tissues in which the  
CC corresponding kinase is expressed. The kinase and antibodies against it  
CC can be used in pharmacogenomic analysis. The kinase can be used to treat  
CC a disorder characterised by an absence of, inappropriate, or unwanted  
CC expression of the kinase. The kinase, homologous to SNF kinases (not  
CC defined) is expressed by a gene located on human chromosome 3. The  
CC present sequence is the cDNA encoding the kinase  
XX  
SQ Sequence 2869 BP; 965 A; 570 C; 645 G; 689 T; 0 U; 0 Other;

Query Match 96.4%; Score 1397.4; DB 6; Length 2869;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1398; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGCTGAAATCCAAAGAGGAGGAGCTAAGTGTGATGGATCAACAGCACTTCCACTTAT 60  
Db 181 ATGCTGAAATCCAAAGAGGAGGAGCTAAGTGTGATGGATCAACAGCACTTCCACTTAT 240  
Qy 61 CCAAGACCTTGATTGCAAGAAGATAGTCTTCAACAAAACTTGCACTGGAAGTTT 120  
Db 241 CCAAGACCTTGATTGCAAGAAGATAGTCTTCAACAAAACTTGCACTGGAAGTTT 300  
Qy 121 GGAACCTGTCTATCTGTTTCAGACAAGAAAGCCAAACGAGGAGGAAATTAAGGTACTT 180  
Db 301 GGAACCTGTCTATCTGTTTCAGACAAGAAAGCCAAACGAGGAGGAAATTAAGGTACTT 360  
Qy 181 AAGGAATATCTGTTGAGAACTAAATCCAAATGAAACTGTACAGGCAATTTGGAAGCC 240  
Db 361 AAGGAATATCTGTTGAGAACTAAATCCAAATGAAACTGTACAGGCAATTTGGAAGCC 420  
Qy 241 CAACCTCTCCAGCTGGACCCAGCCACTTGTCAAGTTCCATGCAAGTTTGTGGAG 300  
Db 421 CAACCTCTCCAGCTGGACCCAGCCACTTGTCAAGTTCCATGCAAGTTTGTGGAG 480  
Qy 301 CAAGATAATTTCTGCATTATCAGCGAGTACTGTGAGGCGCGAGATCTGGACGATAAAAT 360  
Db 481 CAAGATAATTTCTGCATTATCAGCGAGTACTGTGAGGCGCGAGATCTGGACGATAAAAT 540  
Qy 361 CAGGAATATAACAGCTGGAAAATCTTTCCAGAAAATCAAAATATAGATGTTTATC 420  
Db 541 CAGGAATATAACAGCTGGAAAATCTTTCCAGAAAATCAAAATATAGATGTTTATC 600  
Qy 421 CAGCTGCTGTGGAGTTGACTACATGTCATGAGAGGAGGATCTTCATCGAGACTTAAAG 480  
Db 601 CAGCTGCTGTGGAGTTGACTACATGTCATGAGAGGAGGATCTTCATCGAGACTTAAAG 660  
Qy 481 TCAAGAAATGTAATTTCTGAAAATATATCTCTTAAATTTGGAGATTTTGGAGTTTCTCGA 540



241 CAACTCTCTCCAAAGCTGGACCCAGCCAGCCATTGTCAAGTTCCATGCAAGTTTGTGGAG 300  
Db CAACTCTCTCCAAAGCTGGACCCAGCCAGCCATTGTCAAGTTCCATGCAAG-TTGTGGAG 452  
301 CAGCATATTTCTGCATTATCAGCGAGTACTGTGAGGCGCGAGATCTGGACGATAAAATT 360  
Db CAGCATATTTCTGCATTATCAGCGAGTACTGTGAGGCGCGAGATCTGGACGATAAAATT 512  
361 CAGGAATATAAACAAGCTGGAAAAATCTTTCCAGAAAAATCAAAATAATAGAAATGTTTATC 420  
Db CAGGAATATAAACAAGCTGGAAAAATCTTTCCAGAAAAATCAAAATAATAGAAATGTTTATC 572  
421 CAGCTGCTGCTGGAGTTGACTACATCATGAGAGGAGGATCTTATCGAGACTTAAAG 480  
Db CAGCTGCTGCTGGAGTTGACTACATCATGAGAGGAGGATCTTATCGAGACTTAAAG 632  
481 TCAAGAATGTATTTCTGAAAAATTAATCTCTTAAAAATTCGAGATTTTGGAGTTTCTCGA 540  
Db TCAAGAATGTATTTCTGAAAAATTAATCTCTTAAAAATTCGAGATTTTGGAGTTTCTCGA 592  
541 CTTCATATGGATCTGTGACCTGGCCACAACTTTAACTGGAACCTCCCATATATAGAT 600  
Db CTTCATATGGATCTGTGACCTGGCCACAACTTTAACTGGAACCTCCCATATATAGAT 752  
601 CTTGAGGCTCTGAAACACCAAGGCTATGACACAAAGTCGAGACATCTGTCACCTGGCATGC 660  
Db CTTGAGGCTCTGAAACACCAAGGCTATGACACAAAGTCGAGACATCTGTCACCTGGCATGC 812  
661 ATTTGATGAGATGTGCTCCATGAATCAATGCAATCTGCTGGCTCCAAATTTCTTATCCATT 720  
Db ATTTGATGAGATGTGCTCCATGAATCAATGCAATCTGCTGGCTCCAAATTTCTTATCCATT 872  
721 GTTTTAAAAATTTCTGAAGTGACACACCTTCTCTCCCTGAGAGATATCAAAAGAACTA 780  
Db GTTTTAAAAATTTCTGAAGTGACACACCTTCTCTCCCTGAGAGATATCAAAAGAACTA 932  
781 AATGCCATCATGAAAGCATGTTGAAACAAGATCTTCAATTAAGACCATCTGCTATCGAA 840  
Db AATGCCATCATGAAAGCATGTTGAAACAAGATCTTCAATTAAGACCATCTGCTATCGAA 992  
841 ATTTTAAAAATTCCTTACCTTGATGAGAGCTACAGAACCTTAATGTGTAGATATTAGAA 900  
Db ATTTTAAAAATTCCTTACCTTGATGAGAGCTACAGAACCTTAATGTGTAGATATTAGAA 1052  
901 ATGACTCTGGAAGACAAAAATTTGGATTTGCAAGAGGCTGCTCATATATTAATGCTC 960  
Db ATGACTCTGGAAGACAAAAATTTGGATTTGCAAGAGGCTGCTCATATATTAATGCTC 1112  
961 ATGCAAAAAAGGATCCACTGAGACTCTGAGGGCACTCTCAGAACTACAGAAATGACG 1020  
Db ATGCAAAAAAGGATCCACTGAGACTCTGAGGGCACTCTCAGAACTACAGAAATGACG 1172  
1021 CCAAGAAAGGATGCGCTGAGGAGCTCCAGGGGCTGTGAGAAAGCCAGGAGCTG 1080  
Db CCAAGAAAGGATGCGCTGAGGAGCTCCAGGGGCTGTGAGAAAGCCAGGAGCTG 1232  
1081 AAAAGATTGTGAAGAAAAATATGAAGAAAAATAGCAAAACGAATGCAAGAAATTTGAGATCT 1140  
Db AAAAGATTGTGAAGAAAAATATGAAGAAAAATAGCAAAACGAATGCAAGAAATTTGAGATCT 1292  
1141 CGGAACTTTACAGCTGAGTGTGATGTACTTCCATGAAACACATTTAAAGGAATG 1200  
Db CGGAACTTTACAGCTGAGTGTGATGTACTTCCATGAAACACATTTAAAGGAATG 1352  
1201 GAAGAAAGGAGGACCACTGAGGGAAGACTTTCTTGTTCACCCAGGAGGAGATGAA 1260  
Db GAAGAAAGGAGGACCACTGAGGGAAGACTTTCTTGTTCACCCAGGAGGAGATGAA 1412  
1261 GAGAGTGCACAGGAGGAGGAAATCTGATGACCACTTTAGAGAACTGCTGCTGAG 1320  
Db GAGAGTGCACAGGAGGAGGAAATCTGATGACCACTTTAGAGAACTGCTGCTGAG 1472

1321 TCTCAGCCTATTCTTCCATGAGCCTCCAGAACTTGAATCAATTTGTAGAGATGCCACA 1380  
Db TCTCAGCCTATTCTTCCATGAGCCTCCAGAACTTGAATCAATTTGTAGAGATGCCACA 1532  
1381 TCTGACCTTGGATACCATG 1399  
Db TCTGACCTTGGATACCATG 1551

RESULT 7  
ADC99143  
ID ADC99143 standard; cdna; 2688 BP.  
XX  
AC ADC99143;  
XX  
DT 01-JAN-2004 (first entry)  
XX Human KPP cdna - SEQ ID 96.  
DE  
XX  
KW anti-HIV; anti-allergic; anti-inflammatory; antianemic; antiparkinsonian;  
nootropic; anticonvulsant; antiarteriosclerotic; antiasthmatic;  
immunosuppressive; antithyroid; cytostatic; hepatotropic; dermatological;  
antidiabetic; nephrotropic; antigout; thyromimetic; neuroprotective;  
osteopathic; antichratic; antiparasitic; antihelminthic; antipsoriatic;  
uropathic; ophthalmological; antirheumatic; haemostatic; antibacterial;  
vicide; protozoicide; fungicide; kinase; phosphatase; KPP;  
cell proliferative disorder; atherosclerosis; cirrhosis; hepatitis;  
cancer; developmental; mental retardation; neurological;  
Alzheimer's disease; Parkinson's; autoimmune; inflammatory;  
diabetes mellitus; viral; bacterial; fungal; parasitic; protozoan;  
helminthic infection; transgenic; gene therapy; human; ss; gene.

OS Homo sapiens.  
XX  
XX WO2003033680-A2.  
XX  
PD 24-APR-2003.  
XX  
XX 17-OCT-2002; 2002WO-US033723.  
XX  
PF 19-OCT-2001; 2001US-0345474P.  
XX  
PR 02-NOV-2001; 2001US-0343910P.  
PR 13-NOV-2001; 2001US-0333098P.  
PR 16-NOV-2001; 2001US-0332424P.  
PR 30-NOV-2001; 2001US-0334288P.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Bandman O, Baughn MR, Becha SD, Borowsky ML, Duggan BM;  
Emerling BM, Forsythe IJ, Gandhi AR, Gorvad AB, Griffin JA;  
Gururajan R, Hafalia AJA, Khan FA, Lal PG, Lee BA, Lee SY;  
Lindquist EA, Lu DAM, Lu Y, Marquis JP, Nguyen DB, Arvizu CS;  
Rankumar J, Recipon SA, Richardson TW, Swainakar A, Tang YT;  
Thornton MB, Tran UK, Chawla NK, Warren BA, Yang J, Rao MG, Yue H;  
Zebbarjadian Y;  
XX  
XX WPI; 2003-403214/38.  
XX  
XX P-PSDB; ADC99091.  
XX  
XX New human kinases and phosphatases and polynucleotides, useful for  
diagnosing, treating or preventing autoimmune or inflammatory disorders  
(e.g. AIDS, allergy or anemia), multiple sclerosis, osteoarthritis,  
cancer or hepatitis.  
XX  
XX Claim 5; SEQ ID NO 96; 424pp; English.  
XX  
XX The invention relates to a novel isolated polypeptide which is a human  
kinase and phosphatase (KPP). The KPP polypeptides, polynucleotides,  
agonists and antagonists are useful for diagnosing, treating or  
preventing cell proliferative disorders such as atherosclerosis,  
cirrhosis, hepatitis and cancer, developmental disorders e.g. mental  
retardation, neurological disorders including Alzheimer's disease and  
Parkinson's disease, autoimmune and inflammatory disorders such as

CC Crohn's disease and diabetes mellitus and finally, viral, bacterial, fungal, parasitic, protozoan or helminthic infections. Furthermore, the polynucleotides encoding KPP may be useful for creating transgenic animals to model human disease, as well as during gene therapy procedures. The current sequence is that of the human KPP cDNA of the CC invention.

XX  
SQ Sequence 2688 BP; 842 A; 562 C; 615 G; 669 T; 0 U; 0 Other;

Query Match 81.4%; Score 1179; DB 9; Length 2688;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTGAAATTCAGAGGAGCAGCTAAGTGTGTGAGTGTGATCAACAGCCATTTCCTTAT 60  
DB 251 ATGCTGAAATTCAGAGGAGCAGCTAAGTGTGTGAGTGTGATCAACAGCCATTTCCTTAT 310  
QY 61 CCAAGACCTTGTTCAGAGAGATACGCTTCAACAAACTTGGCAGTGGAGTTT 120  
DB 311 CCAAGACCTTGTTCAGAGAGATACGCTTCAACAAACTTGGCAGTGGAGTTT 370  
QY 121 GGAAGTGTCTATCTGTTTCAGACAAAGAAAGCCAAACGAGGAGAGAAATTAAGGTACTT 180  
DB 371 GGAAGTGTCTATCTGTTTCAGACAAAGAAAGCCAAACGAGGAGAGAAATTAAGGTACTT 430  
QY 181 AAGGAATATCTTGGAGAACCTAATCCAAATGAATCTGTACAGCCAAATTTGAGGCC 240  
DB 431 AAGGAATATCTTGGAGAACCTAATCCAAATGAATCTGTACAGCCAAATTTGAGGCC 490  
QY 241 CAATCTCTCCAAAGTGGACCAACCCAGGACCTTCAAGTTCATCAAGTTTCTGAGGAG 300  
DB 491 CAATCTCTCCAAAGTGGACCAACCCAGGACCTTCAAGTTCATCAAGTTTCTGAGGAG 550  
QY 301 CAAGATAATTTCTGCATTATCACGGAGTACTGTGAGGCGCGAGATCTGCAGCATTAAG 360  
DB 551 CAAGATAATTTCTGCATTATCACGGAGTACTGTGAGGCGCGAGATCTGCAGCATTAAG 610  
QY 361 CAGGATATAAACAGCTGGAAAATTTTCCAGAAAATCAATATAGATGGTTTATC 420  
DB 611 CAGGATATAAACAGCTGGAAAATTTTCCAGAAAATCAATATAGATGGTTTATC 670  
QY 421 CAGCTGCTGCTGGAGTTGACTACATGCTAGAGGAGGATCTTATCGAGCATTAAG 480  
DB 671 CAGCTGCTGCTGGAGTTGACTACATGCTAGAGGAGGATCTTATCGAGCATTAAG 730  
QY 481 TCAAGAAATGTATTTCTGAAAATAATCTCTTAAATTTGGAGATTTGGAGTTTCTCGA 540  
DB 731 TCAAGAAATGTATTTCTGAAAATAATCTCTTAAATTTGGAGATTTGGAGTTTCTCGA 790  
QY 541 CTCTTAATGGGATCTGTGACCTGGCCACAACTTTAACTGGAATCTCCCATTTATATGAT 600  
DB 791 CTCTTAATGGGATCTGTGACCTGGCCACAACTTTAACTGGAATCTCCCATTTATATGAT 850  
QY 601 CCTGAGGCTCTGAAACCAAGGCTATGACACAAAGTCGGACATCTGGTCACTGGCATGC 660  
DB 851 CCTGAGGCTCTGAAACCAAGGCTATGACACAAAGTCGGACATCTGGTCACTGGCATGC 910  
QY 661 ATTTTGTATGAGATGCTGCTGATGAATCATGATTCGCTGGCTCCCAATTTCTTATCCAT 720  
DB 911 ATTTTGTATGAGATGCTGCTGATGAATCATGATTCGCTGGCTCCCAATTTCTTATCCAT 970  
QY 721 GTTTTAAATTTGTTGAAGTGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTA 780  
DB 971 GTTTTAAATTTGTTGAAGTGTGACACACCTTCTCTCCCTGAGAGATATCCAAAGAACTA 1030  
QY 781 AATGCCATCATGGAAGAGATGTTGAACAAGATCCTTCATTAAGACCATCTGCTATCGAA 840  
DB 1031 AATGCCATCATGGAAGAGATGTTGAACAAGATCCTTCATTAAGACCATCTGCTATCGAA 1090  
QY 841 ATTTTAAATTCCTTACCTTGTATGAGCAGCTACAGAACCTAATGTGTAGATATTCAGAA 900  
DB 1091 ATTTTAAATTCCTTACCTTGTATGAGCAGCTACAGAACCTAATGTGTAGATATTCAGAA 1150

QY 901 ATGACTCTGGAAGACAAAATTTGGATTGTCAAGAGGAGGCTGTCTATATAATTATGTC 960  
DB 1151 ATGACTCTGGAAGACAAAATTTGGATTGTCAAGAGGAGGCTGTCTATATAATTATGTC 1210  
QY 961 ATGCAAAAAGGATCCACCTCGACACTCTGAGGCGACTGTGAGAGTACAGAAAATGACG 1020  
DB 1211 ATGCAAAAAGGATCCACCTCGACACTCTGAGGCGACTGTGAGAGTACAGAAAATGACG 1270  
QY 1021 CCAAGAGAAAGGATCGCGCTGAGGAAGCTCCAGCGGCTGATGAGAAGCCAGGAAGCTG 1080  
DB 1271 CCAAGAGAAAGGATCGCGCTGAGGAAGCTCCAGCGGCTGATGAGAAGCCAGGAAGCTG 1330  
QY 1081 AAAAGAGTTTGGAGAAATATGAGAAATATGAGAAATATGAGAAATATGAGAAATATGAG 1140  
DB 1331 AAAAGAGTTTGGAGAAATATGAGAAATATGAGAAATATGAGAAATATGAGAAATATGAG 1390  
QY 1141 CGGAACCTTTCAGCAGCTGAGTGTGATGTACTCCATGAA 1179  
DB 1391 CGGAACCTTTCAGCAGCTGAGTGTGATGTACTCCATGAA 1429  
RESULT 8  
ADE28343  
ID ADE28343 standard; cDNA; 2631 BP.  
XX  
AC ADE28343;  
XX 29-JAN-2004 (first entry)  
XX Human KPP cDNA - SEQ ID 54.  
XX kinase; phosphatase; KPP; hepatotropic; antiarteriosclerotic;  
KW antiproliferative; cytosolic; haemostatic; muscular; cerebroprotective;  
KW neotropic; ophthalmological; anticonvulsant; vasotropic; neuroprotective;  
KW antiparkinsonian; antiasthmatic; antianemic; antiasthmatic;  
KW antidiabetic; antiinflammatory; osteopathic; antibacterial;  
KW antirheumatic; dermatological; virucide; antibacterial; fungicide;  
KW antiparasitic; protozoic; antihelminthic; antitumor; cardiovascular;  
KW antiarteriosclerotic; immunosuppressive; cell proliferative; cirrhosis;  
KW hepatitis; arteriosclerosis; psoriasis; primary thrombocytopenia; cancer;  
KW developmental; renal tubular acidosis; Becker's muscular dystrophy;  
KW gonadal dysgenesis; hypothyroidism; seizure; neurological;  
KW Pick's disease; cataract; epilepsy; ischaemic cerebrovascular; stroke;  
KW Alzheimer's; Parkinson's; dementia; autoimmune; inflammatory; AIDS;  
KW allergy; anaemia; asthma; diabetes mellitus; bronchitis; osteoporosis;  
KW osteoarthritis; rheumatoid arthritis; contact dermatitis; gout;  
KW lipid disorder; cholestasis; Gaucher's; diabetes; atherosclerosis; liver;  
KW viral; bacterial; fungal; parasitic; protozoan; helminthic infection;  
KW trauma; gene therapy; human; ss; gene.  
XX Homo sapiens.  
OS  
XX WO2003080805-A2.  
PN  
XX 02-OCT-2003.  
PD  
XX 18-MAR-2003; 2003WO-US008715.  
PF  
XX 19-MAR-2002; 2002US-0366089P.  
PR 29-MAR-2002; 2002US-0369248P.  
XX (INCY) INCYTE CORP.  
PA  
XX Chien D, Jin P, Hawkins PR, Baughn MR, Becha SD, Chang H;  
PI Ding L, Elliott VS, Emerling BM, Gandhi AR, Gietzen KJ, Griffin JA;  
PI Gururajan R, Hafalia AUA, Ison CH, Kable AE, Khare R, Lee SY;  
PI Lee EA, Lu Y, Marquis JP, Lehr-Wason PM, Ramkumar J, Richardson TW;  
PI Swarnakar A, Tran UK, Chawla NK, Yao MG, Yue H, Bhatia U;  
PI Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;  
XX WPI; 2004-011523/01.  
DR P-PSDB; ADE28291.  
XX



PT New human kinases and phosphatases, and polynucleotides encoding them,  
PT useful for treating, preventing or diagnosing e.g. cell proliferative  
PT disorders, inflammatory, autoimmune, viral, bacterial, parasitic or  
XX fungal diseases.  
PS Claim 5; SEQ ID NO 54; 340pp; English.  
XX  
CC The invention relates to a novel isolated kinase and phosphatase (KPP)  
CC polypeptide. The polypeptide of the invention demonstrates hepatotropic,  
CC antiarteriosclerotic, antiparasitic, cytostatic, haemostatic, muscular,  
CC cerebroprotective, neurotropic, ophthalmological, anticonvulsant,  
CC vasotrophic, neuroprotective, antiparkinsonian, antidiabetic,  
CC antianemic, antirheumatic, dermatological, antihelminthic, osteopathic,  
CC antiarthritic, antiparasitic, antidiabetic, antihelminthic, osteopathic,  
CC fungicide, antiparasitic, protozoacide, antihelminthic, antitubercular,  
CC cardiovascular, antiarteriosclerotic and immunosuppressive activities.  
CC The KPP polypeptides may be useful for diagnosing, treating or preventing  
CC cell proliferative disorders including cirrhosis, hepatitis,  
CC arteriosclerosis, psoriasis, primary thrombocytopenia and cancer,  
CC developmental disorders such as renal tubular acidosis, Becker's muscular  
CC dystrophy, gonadal dysgenesis, hypothyroidism or seizures, neurological  
CC disorders e.g. Pick's disease, cataract, epilepsy, ischaemic  
CC cerebrovascular disease, stroke, Alzheimer's disease, Parkinson's disease  
CC or dementia, autoimmune or inflammatory disorders including AIDS,  
CC allergies, anaemia, asthma, diabetes mellitus, bronchitis, osteoporosis,  
CC osteoarthritis, rheumatoid arthritis, contact dermatitis or gout and  
CC lipid disorders such as cholestasis, Gaucher's disease, diabetes,  
CC atherosclerosis or liver disease, as well as viral, bacterial, fungal,  
CC parasitic, protozoan or helminthic infections and trauma. Furthermore,  
CC the polypeptide may be utilised during gene therapy procedures. The  
CC current sequence is that of the human KPP cDNA of the invention.  
XX  
SQ Sequence 2631 BP; 833 A; 539 C; 619 G; 640 T; 0 U; 0 Other;

Query Match 73.48; Score 1063; DB 10; Length 2631;  
Best Local Similarity 100.0%; Pred. No. 2.9e-300;  
Matches 1063; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 337 GCGGAGATCTGGACGATAAAATTCAGGATATAAACAAGCTGGAATAATCTTCCAGAA 396  
DB 421 GCGGAGATCTGGACGATAAAATTCAGGATATAAACAAGCTGGAATAATCTTCCAGAA 480  
QY 397 AATCAATATAGATGTTTATCCAGCTCTCTCGGAGTTGACATGATGATGAGG 456  
DB 481 AATCAATATAGATGTTTATCCAGCTCTCTCGGAGTTGACATGATGATGAGG 540  
QY 457 AGGATATCTTCATCGAGACTTAAAGTCAAGATGTTTCTGAAAAATATCTCTTAA 516  
DB 541 AGGATATCTTCATCGAGACTTAAAGTCAAGATGTTTCTGAAAAATATCTCTTAA 600  
QY 517 ATTGGAGATTTGGAGTTTCGACTTCTAATGGGATCTGTGACCTGGCCCAACTTTA 576  
DB 601 ATTGGAGATTTGGAGTTTCGACTTCTAATGGGATCTGTGACCTGGCCCAACTTTA 660  
QY 577 ACTGGAACTCCCATATATAGTCTCTGAGGCTCTGAAACACCAAGCTATGACACAAAG 636  
DB 661 ACTGGAACTCCCATATATAGTCTCTGAGGCTCTGAAACACCAAGCTATGACACAAAG 720  
QY 637 TCGGACATCTGGTCACTGGCATGCAATTTGTATGAGATGTGTGATGATGATGATGATG 696  
DB 721 TCGGACATCTGGTCACTGGCATGCAATTTGTATGAGATGTGTGATGATGATGATGATG 780  
QY 697 GCTGGCTCCAAATTTCTATCATGTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 756  
DB 781 GCTGGCTCCAAATTTCTATCATGTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 840  
QY 757 CTTGAGAGATATCCAAAGAACTAAATGTCATCATGGAAGCATGTTTGAACAGAAATCT 816  
DB 841 CTTGAGAGATATCCAAAGAACTAAATGTCATCATGGAAGCATGTTTGAACAGAAATCT 900  
QY 817 TCATTAAAGCACTCTGCTATCGAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 876  
DB 901 TCATTAAAGCACTCTGCTATCGAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAAT 960

QY 877 AACCTAATGTGTAGATATTCAGAAATGACTCTGGAAGACAAAATTTTGGATTGTGAGAAG 936  
DB 961 AACCTAATGTGTAGATATTCAGAAATGACTCTGGAAGACAAAATTTTGGATTGTGAGAAG 1020  
QY 937 GAGGCTGTCTATATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 996  
DB 1021 GAGGCTGTCTATATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 1080  
QY 997 CTGTCTAGAAAGTACAGAAATGACGCAAGAAAGGATGCGGCTGAGGAAGCTCCAGGCG 1056  
DB 1081 CTGTCTAGAAAGTACAGAAATGACGCAAGAAAGGATGCGGCTGAGGAAGCTCCAGGCG 1140  
QY 1057 GCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1116  
DB 1141 GCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200  
QY 1117 AAACGAATGCAAGATTCAGATCTCGGAATCTTCAGCAGCTGAGTGTGATGATGATGATGAT 1176  
DB 1201 AAACGAATGCAAGATTCAGATCTCGGAATCTTCAGCAGCTGAGTGTGATGATGATGATGAT 1260  
QY 1177 GAAATACACATTTAAAGGAATGGAAGAAAGGAGGAACTCTGAGGGAAGACTTTCT 1236  
DB 1261 GAAATACACATTTAAAGGAATGGAAGAAAGGAGGAACTCTGAGGGAAGACTTTCT 1320  
QY 1237 TGTTCACCCAGGACGAGATGGAAGAGGTGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 1296  
DB 1321 TGTTCACCCAGGACGAGATGGAAGAGGTGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 1380  
QY 1297 CCAATTTAGAGAACCTGCTGAGTCTCAGCCTATTCCTTCCATGAGGACCTCCAGAACTT 1356  
DB 1381 CCAATTTAGAGAACCTGCTGAGTCTCAGCCTATTCCTTCCATGAGGACCTCCAGAACTT 1440  
QY 1357 GAATCAATTTAGAGGATGCCACATCTGACCTTGATACCATG 1399  
DB 1441 GAATCAATTTAGAGGATGCCACATCTGACCTTGATACCATG 1483

RESULT 9  
ADC99142  
ID ADC99142 standard; cDNA; 2483 BP.  
XX  
AC ADC99142;  
XX  
DT 01-JAN-2004 (first entry)  
XX  
DE Human KPP cDNA - SEQ ID 95.  
XX  
KW anti-HIV; anti-allergic; anti-inflammatory; antianemic; antiparkinsonian;  
KW neurotropic; anticonvulsant; antiarteriosclerotic; antidiabetic;  
KW immunosuppressive; antihypertensive; cytostatic; hepatotropic; dermatological;  
KW antidiabetic; nephrotropic; antitubercular; thyromimetic; neuroprotective;  
KW osteopathic; antiparasitic; antihelminthic; antiparasitic; antiparasitic;  
KW uropathic; ophthalmological; antirheumatic; haemostatic; antiparasitic;  
KW virucide; protozoacide; fungicide; kinase; phosphatase; KPP;  
KW cell proliferative disorder; atherosclerosis; cirrhosis; hepatitis;  
KW cancer; developmental disorder; mental retardation; neurological;  
KW Alzheimer's disease; Parkinson's; autoimmune; inflammatory; Crohn's;  
KW diabetes mellitus; viral; bacterial; fungal; parasitic; protozoan;  
KW helminthic infection; transgenic; gene therapy; human; ss; gene.  
OS Homo sapiens.  
XX  
XX WO200303680-A2.  
XX  
PD 24-APR-2003.  
XX  
XX 17-OCT-2002; 2002WO-US033723.  
XX  
XX 19-OCT-2001; 2001US-0345474P.  
XX 02-NOV-2001; 2001US-0343910P.  
XX 13-NOV-2001; 2001US-0333098P.  
XX 16-NOV-2001; 2001US-0332424P.



```
PR 30-NOV-2001; 2001US-0334288P.
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Bandman O, Baughn MR, Becha SD, Borowsky ML, Duggan BM;
XX Emerling BM, Forsythe IJ, Gandhi AR, Gorvad AE, Griffin JA;
XX Gururajan R, Hafalia AJA, Khan FA, Lal PG, Lee EA, Lee SY;
XX Lindquist EA, Lu DAM, Lu Y, Marquis JP, Nguyen DB, Arvizu CS;
XX Ramkumar J, Recipon SA, Richardson TW, Swarnakar A, Tang YT;
XX Thornton MB, Tran UK, Chawla NK, Warren BA, Yang J, Yac MG, Yue H;
XX Zebbarjadian Y;
XX WPI: 2003-403214/38.
XX P-PSDB; ADC99090.
XX
XX New human kinases and phosphatases and polynucleotides, useful for
XX diagnosing, treating or preventing autoimmune or inflammatory disorders
XX (e.g. AIDS, allergy or anemia), multiple sclerosis, osteoarthritis,
XX cancer or hepatitis.
XX
XX Claim 5; SEQ ID NO 95; 424pp; English.
XX
XX The invention relates to a novel isolated polypeptide which is a human
XX kinase and phosphatase (KPP). The KPP polypeptides, polynucleotides,
XX agonists and antagonists are useful for diagnosing, treating or
XX preventing cell proliferative disorders such as atherosclerosis,
XX cirrhosis, hepatitis and cancer, developmental disorders e.g. mental
XX retardation, neurological disorders including Alzheimer's disease and
XX Parkinson's disease, autoimmune and inflammatory disorders such as
XX Crohn's disease and diabetes mellitus and finally, viral, bacterial,
XX fungal, parasitic, protozoan or helminthic infections. Furthermore, the
XX polynucleotides encoding KPP may be useful for creating transgenic
XX animals to model human disease, as well as during gene therapy
XX procedures. The current sequence is that of the human KPP cDNA of the
XX invention.
XX
XX SQ Sequence 2483 BP; 783 A; 515 C; 597 G; 588 T; 0 U; 0 Other;
XX
XX Query Match 52.4%; Score 759; DB 9; Length 2483;
XX Best Local Similarity 77.5%; Pred. No. 3.3e-211;
XX Matches 1084; Conservative 0; Mismatches 0; Indels 315; Gaps 1;
XX
XX 1 ATGCTGAAATCCAAAGGAGGAGCTTAAGTGTGTGAGTGGATCAACAGCCATTTCACCTAT 60
XX 251 ATGCTGAAATCCAAAGGAGGAGCTTAAGTGTGTGAGTGGATCAACAGCCATTTCACCTAT 310
XX
XX 61 CCAAGACCTTGTGAGAGAGATAGCTGCTTCAACAAACTTGGCAGTGGAGTTT 120
XX 311 CCAAGACCTTGTGAGAGAGATAGCTGCTTCAACAAACTTGGCAGTGGAGTTT 370
XX
XX 121 GGAATCTGTATCTGTTTTCAGACAAAGAACCCAAAGCAGGAGGAGGAATTAAGGTACTT 180
XX 371 GGAATCTGTATCTGTTTTCAGACAAAGAACCCAAAGCAGGAGGAGGAATTAAGGTACTT 430
XX
XX 181 AAGGAATATCTGTTGAGAACTAAATCCAAATCAAACTGTACAGGCCAATTTGGAGCC 240
XX 431 AAGGAATATCTGTTGAGAACTAAATCCAAATCAAACTGTACAGGCCAATTTGGAGCC 490
XX
XX 241 CAATCTCTCCAAAGCTGGACCCAGCCAGCTTCAAGTTCATGCAAGTTTGTGGAG 300
XX 491 CAATCTCTCCAAAGCTGGACCCAGCCAGCTTCAAGTTCATGCAAGTTTGTGGAG 550
XX
XX 301 CAAGATAATTTCTGCATTATCAAGGATCTGTGAGGGCCGAGATCTGGACGATAAATT 360
XX 551 CAAGATAATTTCTGCATTATCAAGGATCTGTGAGGGCCGAGATCTGGACGATAAATT 610
XX
XX 361 CAGGAATATAACAGCTGGAAATCTTTCCAGAAATCAATAATAGATGTTTATC 420
XX 611 CAGGAATATAACAGCTGGAAATCTTTCCAGAAATCAATAATAGATGTTTATC 670
XX
XX 421 CAGCTGTCTGGAGTTGACTACATCATGATGAGAGGAGGATCTTCAATCGAGACTTAAG 480
XX 671 CAGCTGTCTGGAGTTGACTACATCATGATGAGAGGAGGATCTTCAATCGAGACTTAAG 730
```

RESULT 10  
ABL34984  
ID ABL34984 standard; cDNA; 2538 BP.  
XX  
XX ABL34984;  
XX AC  
XX XX  
DT 04-APR-2002 (first entry)

XX DE Rat cDNA isolated from skin cells SEQ ID NO: 511.  
 XX KW Human; rat; mouse; skin cell; skin wound; cancer; growth defect;  
 XX KW developmental defect; inflammatory disease; dermatological; vulnary;  
 XX KW immunomodulator; anti-inflammatory; cytostatic; neuroprotective; gene;  
 XX KW ss.  
 XX OS Rattus sp.  
 XX PN WO2001:90357-A1.  
 XX XX 29-NOV-2001.  
 XX PD 24-MAY-2001; 2001WO-NZ000099.  
 XX XX 24-MAY-2000; 2000US-0206650P.  
 XX PR 25-JUL-2000; 2000US-0221232P.  
 XX XX (GENE-) GENESIS RES & DEV CORP LTD.  
 XX PA Watson JD, Strachan L, Sleeman M, Onrust R, Murison JG;  
 XX PI Kumble KD;  
 XX PI WPI; 2002-122020/16.  
 XX DR New polynucleotides and polypeptides encoded by the polynucleotides  
 XX PT isolated from skin cells, useful for treating skin wounds, cancers,  
 XX PT growth and developmental defects, inflammatory diseases, or for  
 XX PT modulating immune responses.  
 XX PS Claim 1; Page 316-317; 466pp; English.  
 XX CC The present invention provides the protein and coding sequences of cDNAs  
 XX CC isolated from human, murine and rat skin cell libraries. The sequences  
 XX CC can be used in the development of therapeutic agents useful in the  
 XX CC treatment of skin diseases, including skin wounds, cancer, growth  
 XX CC defects, developmental defects and inflammatory diseases. The proteins  
 XX CC have important roles in the induction of hair growth, cell proliferation  
 XX CC and cell-cell interaction, in maintaining tissue integrity, in wound  
 XX CC healing and in modulating immune responses. The present sequence is a  
 XX CC cDNA of the invention  
 XX CC  
 XX SQ Sequence 2538 BP; 697 A; 558 C; 636 G; 647 T; 0 U; 0 Other;

Query Match 46.2%; Score 669.6; DB 6; Length 2538;  
 Best Local Similarity 82.2%; Pred. No. 5.1e-185;  
 Matches 797; Conservative 0; Mismatches 154; Indels 9; Gaps 2;

Qy 1 ATGCTGAAATCCAGAGGAGCTAGTGTGTGAGTGGATCAACAGCCATTTCCACTTAT 60  
 Db 457 ATGCTGAAATCCAGAGGAGCTAGTGTGTGAGTGGATCAACAGCCATTTCCACTTAT 510

Qy 61 CCAAGACCTTGATGCAAGAGATAGTGTGTGTGAGTGGATCAACAGCCATTTCCACTTAT 120  
 Db 511 CCACCCACTGTGATGCAAGAGATAGTGTGTGTGAGTGGATCAACAGCCATTTCCACTTAT 570

Qy 121 GGAACCTGTCTATCTGTGTTTCCAGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
 Db 571 GGAACCTGTCTATCTGTGTTTCCAGACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 630

Qy 181 AAGAAATATCTGTGAGAGACTTAATCCAAATGAACCTGACAGCCATTTGGAGCC 240  
 Db 631 AAGAAATATCTGTGAGAGACTTAATCCAAATGAACCTGACAGCCATTTGGAGCC 690

Qy 241 CAACCTCTCTCCAGCTGGACCCAGCCATTTGCAAGTGTCCATGCAAGTGTGTGGAG 300  
 Db 691 CAGCTCTCTCCAGCTGGACCCAGCCATTTGCAAGTGTCCATGCAAGTGTGTGGAG 750

Qy 301 CAGATTAATTTCTGCATATCAAGAGTACTGTGAGGCGCGAGATCTGACGATATAAAT 360  
 Db 751 CAGGATCAATTTGTCATATCAAGAGTACTGTGAGGCGCGAGATCTGACGATATAAAT 810

Qy 361 CAGGAATATAAACAAGCTGGAAAAATCTTTCAGAAAAATCAAAATAAGATGGTTTATC 420  
 Db 811 CAGGAATATAAACAAGCTGGAAAAATCTTTCAGAAAAATCAAAATAAGATGGTTTATC 870

Qy 421 CAGCTGCTGGGAGTTGACTACATGATGAGAGGAGATACCTCATCGAGACTTAAAG 480  
 Db 871 CAGTGTGCTGGGAGTTGATTACATGATGAGAGGAGATACCTCATCGAGACTTAAAG 930

Qy 481 TCAAGAATATATTTCTGAAAAATATCTCTTAAATTTGGAGATTTTGGAGTTTCTCGA 540  
 Db 931 TCAAGAATATATTTCTGAAAAATATCTCTTAAATTTGGAGATTTTGGAGTTTCTCGG 990

Qy 541 CTTCTAATGGGATCTGTGACTGGCCCAACTTTTAAGTGAACCTCCCATATATAGT 600  
 Db 991 CTGCTGATGGGTTTCTGAGCTGGCTACAACTCTAACCGGAGCTCCCATATATAGT 1050

Qy 601 CCTGAGGCTCTGAAACACCAAGGCTATGACACAAAGTCGACATCTGCTCACTGGCATGC 660  
 Db 1051 CCAGAGGCCCTGAGCACCAGGCTGTGATGCCAAGTCTGACATCTGCTCACTGGCATGC 1110

Qy 661 ATTTTGTATGAGATGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 720  
 Db 1111 ATTTTATAGAGATGTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1170

Qy 721 GTTTTAAAAATTTGTAAGGTGACACACCTTCTCTCCCTGAGAGATATCCAAAAGAACTA 780  
 Db 1171 GTTTTGAATATCGTTGAAGGCAACACGCTTCTCTCCCTGAGAGATATCCAAAAGAACTA 1230

Qy 781 AATGCCATCATGGAAGCATGTTGAACAAGAAATCTTCAATTAAGACCATCTGCTATCGAA 840  
 Db 1231 AACATCATCATGTCACGCATGTTGAACAAGAGTCCCTCCCTGAGACCGTCCGCTGCAGAG 1290

Qy 841 ATTTTAAAAATCCCTTACCTTATGATGAGCAGCTACAGAACCTTAATGTGTAGATATTCAGAA 900  
 Db 1291 ATTTTAAAAAGCCCTTATGTAAGAGAGTCCCTTCAAGAGCTGATGATTAACATCCAGAG 1350

Qy 901 ATGACTCTGGAGACAA---AAATTTGGATTTGTCAGAGGAGGCTGCTCATATAATTAAT 957  
 Db 1351 GCGACACTGGAGACAAAGAGGAACTACACCTGTCAGAGGAGGCTGCCATGCAGTTAAC 1410

Qy 958 GCCATGCAAA 967  
 Db 1411 GCCATAAGA 1420

RESULT 11  
 ADCS1181  
 ID ADCS1181 standard; DNA; 1071 BP.  
 XX AC ADCS1181;  
 XX DT 18-DEC-2003 (first entry)  
 XX DE Human cell-cycle related protein coding sequence, SEQ ID 7.  
 XX KW Human; cytostatic; cell-cycle related protein; nuclear export;  
 XX KW nuclear-cytoplasm transport; cytotoxic; cell-cycle control;  
 XX KW immunological disease; neurological disease; cancer; gene; ds.  
 XX OS Homo sapiens.  
 XX PH Location/Qualifiers  
 XX FT 1..1071  
 XX FT /\*tag= a  
 XX FT /partial  
 XX FT /product= "Cell-cycle related protein"  
 XX FT /note= "No start or stop codon given"  
 XX PN JP2003144168-A.  
 XX PD 20-MAY-2003.  
 XX PF 14-NOV-2001; 2001JP-00349158.

XX 14-NOV-2001; 2001JP-00349158.  
XX (KOKU-) KOKURITSU YOBO EISEI KENKYUSHO.  
XX (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.  
XX WPI; 2003-818166/77.  
XX P-PSDB; ADCS1182.  
XX Novel DNA or RNA coding a cell-cycle related protein which has nuclear  
XX export function, useful for screening substance that prevent or treat  
XX cell cycle abnormality diseases e.g. immunological disease.  
XX Claim 15; SEQ ID NO 7; 41pp; Japanese.  
XX The present invention relates to novel cell-cycle related protein such as  
XX NIMA (Never-In Mitosis, gene A)-related protein kinase of Nek 9 and  
XX coding sequences such as a cell-cycle related protein (ADCS1176) having  
XX nuclear export function; cell-cycle related protein (ADCS1178) having  
XX nuclear-cytoplasm transport function; cell-cycle related protein and  
XX (ADCS1180) having nuclear export function; cytotoxic function and  
XX transfer function in the nucleus; and/or cell-cycle related protein  
XX (ADCS1182) having transfer function and cytotoxic function in nucleus,  
XX nuclear export function and nuclear cytoplasm transport function. The  
XX coding sequences for these proteins are given in ADCS1175, ADCS1177,  
XX ADCS1179 and ADCS1181. The sequences of the invention are useful for  
XX screening a substance which promotes or suppresses the transfer function  
XX or cytotoxic function in the nucleus, nuclear-cytoplasm transport  
XX function and/or a cell-cycle control function. The sequences are also  
XX useful for treating or diagnosing an immunological disease, neurological  
XX disease or cancer.  
XX Sequence 1071 BP; 365 A; 211 C; 282 G; 213 T; 0 U; 0 Other;  
Query Match 36.9%; Score 535; DB 9; Length 1071;  
Best Local Similarity 100.0%; Pred. No. 8.7e-146;  
Matches 535; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 865 GAGCAGCTACAGAACCTTAATGTGTAGATATTCAGAAATGCTGGAAGCAAAATTTG 924  
Db 1 GAGCAGCTACAGAACCTTAATGTGTAGATATTCAGAAATGCTGGAAGCAAAATTTG 60  
QY 925 GATTGTGAGAGGAGGCTGCTATATTAATGCCATGCAAAAGGATCCACCTGCAG 984  
Db 61 GATTGTGAGAGGAGGCTGCTATATTAATGCCATGCAAAAGGATCCACCTGCAG 120  
QY 985 ACTCTGAGGCACCTGTGAGAAAGTACAGAAATGAGCCCAAGAGAGGATCGCGCTGAG 1044  
Db 121 ACTCTGAGGCACCTGTGAGAAAGTACAGAAATGAGCCCAAGAGAGGATCGCGCTGAG 180  
QY 1045 AAGCTCCAGCGGCTGATGAGAGCCAGCAAGCTGAAGATTTGGAAGAAATAT 1104  
Db 181 AAGCTCCAGCGGCTGATGAGAGCCAGCAAGCTGAAGATTTGGAAGAAATAT 240  
QY 1105 GAAGAAATAGCAAAACCAATGCAAGATTCGGAATTCGGAATTCGAGCTGAGTGT 1164  
Db 241 GAAGAAATAGCAAAACCAATGCAAGATTCGGAATTCGGAATTCGAGCTGAGTGT 300  
QY 1165 GATGTACTCCTATGAAACACATTTTAAAGGATGGAAGAGGAGGAGCAACCTGAG 1224  
Db 301 GATGTACTCCTATGAAACACATTTTAAAGGATGGAAGAGGAGGAGCAACCTGAG 360  
QY 1225 GGAAGACTTTCTTGTTCACCCAGCAGGATGGAAGAGGATGGCAAGGAGGAGAGAG 1284  
Db 361 GGAAGACTTTCTTGTTCACCCAGCAGGATGGAAGAGGATGGCAAGGAGGAGAGAG 420  
QY 1285 GAATCTGATGAACCAATTTAGAGAACTGCTGAGTCTCAGCTATTCCTTCATGGAC 1344  
Db 421 GAATCTGATGAACCAATTTAGAGAACTGCTGAGTCTCAGCTATTCCTTCATGGAC 480  
QY 1345 CTCACCAACTGGAATCAATTTAGAGAGGATGCCACATCTGATACCATG 1399  
Db 481 CTCACCAACTGGAATCAATTTAGAGAGGATGCCACATCTGATACCATG 535

RESULT 12  
ABK43470  
ID ABK43470 standard; cDNA; 451 BP.  
XX  
AC ABK43470;  
XX  
DT 05-JUN-2002 (first entry)  
XX  
DE DNA encoding novel central nervous system protein #50.  
XX  
KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;  
KW hyperproliferative disorder; neoplasm; cardiovascular disorder;  
KW cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;  
KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;  
KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;  
KW adenocarcinoma; reproductive system disorder; testicular feminisation;  
KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;  
KW respiratory disorder; renal disorder; kidney failure; blood disorder;  
KW myocardial infarction; wound healing; cell proliferation; skin aging;  
KW food additive; food preservative; gene therapy; gene; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO200155318-A2.  
XX  
XX 02-AUG-2001.  
XX  
XX 17-JAN-2001; 2001WO-US001332.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 11-JUL-2000; 2000US-0216800P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225474P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226689P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.



```
Db 122 CCTTGATGAGCAGCTACAGAACCTAATGTGTAGATATTTCAGAAATGACTCTGGAAGACAA 181
Qy 918 AATTTGGATTGTCAGAGAGGCTGCTCATATAATTAATGCCATGCAAAAAGGATCCA 977
Db 182 AATTTGGATTGTCAGAGAGGCTGCTCATATAATTAATGCCATGCAAAAAGGATCCA 241
Qy 978 CCTGAGACTCTGAGGCGACTGTCTGAGAGTACAGAAATGACGCAAGAGAAAGGATGCG 1037
Db 242 CCTGAGACTCTGAGGCGACTGTCTGAGAGTACAGAAATGACGCAAGAGAAAGGATGCG 301
Qy 1038 GCTGAGAGTCTCAGG-CGGCTGATGAGAAAGCCAGGAGCTGAAAAAGATTGTGAAG 1096
Db 302 GCTGAGAGTCTCAGGCGGCTGATGAGAAAGCCAGGAGCTGAAAAAGATTGTGAAG 361
Qy 1097 AAAAAATGAGAAAAATAGCAAAAGCAATGCAAGAAATGAGATCTCGGAACCTTTTCAGCAGC 1156
Db 362 AAAAAATGAGAAAAATAGCAAAAGCAATGCAAGAAATGAGATCTCGGAACCTTTTCAGCAGC 421
Qy 1157 TGAGTGTG-ATGTAATCTCCATGAAAA 1182
Db 422 TGAGTGTGATGTAATCTCCATGAAAA 448

RESULT 13
AAS27206
ID AAS27206 standard; cdna; 430 BP.
XX
AC AAS27206;
XX
DT 07-NOV-2001 (first entry)
XX
cDNA encoding novel signal transduction pathway protein, Seq ID 241.
XX
KW Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW antiinflammatory; anti-Hiv; antibacterial; antiinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS; ss;
KW acquired immune deficiency syndrome.
XX
XX Homo sapiens.
XX
XX WO200154733-A1.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001312.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX
XX 04-FEB-2000; 2000US-0180628P.
XX
XX 24-FEB-2000; 2000US-0184664P.
XX
XX 02-MAR-2000; 2000US-0186350P.
XX
XX 16-MAR-2000; 2000US-0189874P.
XX
XX 17-MAR-2000; 2000US-0190076P.
XX
XX 18-APR-2000; 2000US-0198123P.
XX
XX 19-MAY-2000; 2000US-0205151P.
XX
XX 07-JUN-2000; 2000US-0209467P.
XX
XX 28-JUN-2000; 2000US-0214886P.
XX
XX 30-JUN-2000; 2000US-0215135P.
XX
XX 07-JUL-2000; 2000US-0216647P.
XX
XX 07-JUL-2000; 2000US-0216880P.
XX
XX 11-JUL-2000; 2000US-0217487P.
XX
XX 11-JUL-2000; 2000US-0217496P.
XX
XX 14-JUL-2000; 2000US-0218290P.
XX
XX 26-JUL-2000; 2000US-0220963P.
XX
XX 26-JUL-2000; 2000US-0220964P.
XX
XX 14-AUG-2000; 2000US-0224518P.
XX
XX 14-AUG-2000; 2000US-0224519P.
XX
XX 14-AUG-2000; 2000US-0225213P.
XX
XX 14-AUG-2000; 2000US-0225214P.
XX
XX 14-AUG-2000; 2000US-0225266P.
XX
XX 14-AUG-2000; 2000US-0225267P.
XX
XX 14-AUG-2000; 2000US-0225268P.
XX
XX 14-AUG-2000; 2000US-0225270P.
XX
XX 14-AUG-2000; 2000US-0225447P.
XX
XX 14-AUG-2000; 2000US-0225757P.
XX
XX 14-AUG-2000; 2000US-0225758P.
XX
XX 14-AUG-2000; 2000US-0225759P.
XX
XX 18-AUG-2000; 2000US-0226279P.
XX
XX 22-AUG-2000; 2000US-0226681P.
XX
XX 22-AUG-2000; 2000US-0226688P.
XX
XX 22-AUG-2000; 2000US-0227182P.
XX
XX 23-AUG-2000; 2000US-0227009P.
XX
XX 30-AUG-2000; 2000US-0228924P.
XX
XX 01-SEP-2000; 2000US-0229287P.
XX
XX 01-SEP-2000; 2000US-0229343P.
XX
XX 01-SEP-2000; 2000US-0229344P.
XX
XX 01-SEP-2000; 2000US-0229345P.
XX
XX 05-SEP-2000; 2000US-0229509P.
XX
XX 05-SEP-2000; 2000US-0229513P.
XX
XX 06-SEP-2000; 2000US-0230437P.
XX
XX 06-SEP-2000; 2000US-0230438P.
XX
XX 08-SEP-2000; 2000US-0231242P.
XX
XX 08-SEP-2000; 2000US-0231243P.
XX
XX 08-SEP-2000; 2000US-0231244P.
XX
XX 08-SEP-2000; 2000US-0231413P.
XX
XX 08-SEP-2000; 2000US-0231414P.
XX
XX 08-SEP-2000; 2000US-0232080P.
XX
XX 08-SEP-2000; 2000US-0232081P.
XX
XX 12-SEP-2000; 2000US-0231968P.
XX
XX 14-SEP-2000; 2000US-0232397P.
XX
XX 14-SEP-2000; 2000US-0232398P.
XX
XX 14-SEP-2000; 2000US-0232399P.
XX
XX 14-SEP-2000; 2000US-0232400P.
XX
XX 14-SEP-2000; 2000US-0232401P.
XX
XX 14-SEP-2000; 2000US-0233063P.
XX
XX 14-SEP-2000; 2000US-0233064P.
XX
XX 14-SEP-2000; 2000US-0233065P.
XX
XX 21-SEP-2000; 2000US-0234223P.
XX
XX 21-SEP-2000; 2000US-0234274P.
XX
XX 25-SEP-2000; 2000US-0234997P.
XX
XX 25-SEP-2000; 2000US-0234998P.
XX
XX 26-SEP-2000; 2000US-0235484P.
XX
XX 27-SEP-2000; 2000US-0235834P.
XX
XX 27-SEP-2000; 2000US-0235836P.
XX
XX 29-SEP-2000; 2000US-0236327P.
XX
XX 29-SEP-2000; 2000US-0236367P.
XX
XX 29-SEP-2000; 2000US-0236368P.
XX
XX 29-SEP-2000; 2000US-0236369P.
XX
XX 29-SEP-2000; 2000US-0236370P.
XX
XX 02-OCT-2000; 2000US-0236802P.
XX
XX 02-OCT-2000; 2000US-0237037P.
XX
XX 02-OCT-2000; 2000US-0237038P.
XX
XX 02-OCT-2000; 2000US-0237039P.
XX
XX 02-OCT-2000; 2000US-0237040P.
XX
XX 13-OCT-2000; 2000US-0239935P.
XX
XX 13-OCT-2000; 2000US-0239937P.
XX
XX 20-OCT-2000; 2000US-0240960P.
XX
XX 20-OCT-2000; 2000US-0241221P.
XX
XX 20-OCT-2000; 2000US-0241785P.
XX
XX 20-OCT-2000; 2000US-0241786P.
XX
XX 20-OCT-2000; 2000US-0241787P.
XX
XX 20-OCT-2000; 2000US-0241808P.
XX
XX 20-OCT-2000; 2000US-0241809P.
XX
XX 20-OCT-2000; 2000US-0241826P.
XX
XX 01-NOV-2000; 2000US-0244617P.
XX
XX 08-NOV-2000; 2000US-0246474P.
XX
XX 08-NOV-2000; 2000US-0246475P.
XX
XX 08-NOV-2000; 2000US-0246476P.
XX
XX 08-NOV-2000; 2000US-0246477P.
XX
XX 08-NOV-2000; 2000US-0246478P.
XX
XX 08-NOV-2000; 2000US-0246523P.
```

PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246529P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251888P.  
PR 08-DEC-2000; 2000US-0251889P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254037P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-465460/50.  
XX P-PSDB; AAU17289.  
XX Novel polypeptides useful for diagnosing, treating, preventing and/or  
XX prognosing disorders related to the proteins, including cancers, immune  
XX disorders and neuronal disorders.  
XX Claim 1; SEQ ID NO 241; 880pp; English.  
XX The invention relates to novel isolated polypeptides (I), and  
XX polynucleotides (II). (I), (II) and the antibody to (I) are useful for  
XX diagnosing, preventing and treating diseases including immune system  
XX disorders (e.g. congenital and acquired immunodeficiencies, autoimmune  
XX disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ  
XX transplant rejections and graft versus host disease, infectious diseases  
XX (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and  
XX other blood-related disorders (sickle cell anemia), myeloproliferative  
XX disorders, primary haematopoietic disorders, hyperproliferative disorders  
XX (e.g. Gaucher's disease and cancer), neurodegenerative disorders (e.g.  
XX Alzheimer's disease, Parkinson's disease), chromosomal abnormalities  
XX (Down syndrome), ischaemic injury (e.g. stroke), renal disorders (e.g.  
XX glomerulonephritis), cardiovascular disorders (e.g. arrhythmia),  
XX respiratory disorders, dermatological disorders, in wound healing,  
XX epithelial cell proliferation, endocrine disorders (e.g. Addison's  
XX disease), reproductive system disorders, gastrointestinal disorder  
XX (inflammatory disorders), liver disorders (cirrhosis), as stimulators of

CC B-cell responsiveness to pathogens, activators of T-cells, to induce  
CC higher affinity antibodies, and as a means to induce tumour proliferation  
CC in pathologies e.g. acquired immune deficiency syndrome (AIDS). AAS26976-  
CC AAS27850 represent novel signal transduction pathway protein coding  
CC sequences and PCR primers of the invention  
XX  
Query Match 27.3%; Score 395.4; DB 4; Length 430;  
Best Local Similarity 97.9%; Pred. No. 4.2e-105;  
Matches 418; Conservative 1; Mismatches 6; Indels 2; Gaps 2;  
QY 758 CTGAGAGATATCCAAAAGAACTAAATGCCATCATCGGAAGCATTTGACAGAGATCCTT 817  
DB 1 CTGAGAGATATCCAAAAGAACTAAATGCCATCATCGGAAGCATTTGACAGAGATCCTT 60  
QY 818 CATTAAAGACCATCTGCTATCGAAATTTTAAAAATCCCTTTACCTTTGATGAGCAGCTACAGA 877  
DB 61 CATTAAAGACCATCTGCTATCGAAATTTTAAAAATCCCTTTACCTTTGATGAGCAGCTACAGA 120  
QY 878 ACCTAATGTGTAGATATTTCAGAAATGACTCTGGAAGACAAAATTTGGATTGTTCAGAGG 937  
DB 121 ACCTAATGTGTAGATATTTCAGAAATGACTCTGGAAGACAAAATTTGGATTGTTCAGAGG 180  
QY 938 AGGCTGCTCATATATTAAATGCCATGCAAAAAGATCCACTGCAGACTCTGAGGGCAC 997  
DB 181 AGGCTGCTCATATATTAAATGCCATGCAAAAAGATCCACTGCAGACTCTGAGGGCAC 240  
QY 998 TGTCAAGAGTACAGAAAATGACGCCAAGAGAAAGATCGCGCTGAGGAAGTCCAGG-CG 1056  
DB 241 TGTCAAGAGTACAGAAAATGACGCCAAGAGAAAGATCGCGCTGAGGAAGTCCAGGCGC 300  
QY 1057 GCTGATGAGAAAGCCAGCAAGCTGAAAAGATTGTGGAGAAAATATGAGAAAATAGC 1116  
DB 301 GCTGATGAGAAAGCCAGCAAGCTGAAAAGATTGTGGAGAAAATATGAGAAAATAGC 360  
QY 1117 AAACGAATGCAAGATTCAGATCTCGGAACCTTTTCAGCAGCTGAGTGTG-ATGTACTCCA 1175  
DB 361 AAACGAATGCAAGATTCAGATCTCGGAACCTTTTCAGCAGCTGAGTGTGTAATGTACTTCA 420  
QY 1176 TGA AAAA 1182  
DB 421 TGA AAAA 427  
RESULT 14  
AAK8454  
ID AAK8454 standard; cDNA; 430 BP.  
XX  
AC AAK8454;  
XX  
DT 05-NOV-2001 (first entry)  
XX  
DE Human digestive system antigen coding sequence SEQ ID NO: 770.  
XX  
KW Human; digestive system antigen; gene therapy; cancer; appendicitis;  
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;  
KW digestive system disorder; Meckel's diverticulum; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200155314-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001324.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.



including cancer, Meckel's diverticulum, bacterial or parasitic  
infections, appendicitis, Hirschsprung's disease, chronic colitis or  
ulcerative colitis. The present sequence is a cDNA encoding a digestive  
system antigen of the invention

Sequence 430 BP; 161 A; 76 C; 98 G; 90 T; 0 U; 5 Other;

Query Match 27.3%; Score 395.4; DB 4; Length 430;  
Best Local Similarity 97.9%; Pred. No. 4.2e-105;  
Matches 418; Conservative 1; Mismatches 6; Indels 2; Gaps 2;

QY 758 CTGAGAGATATCCAAAGAACTAAATGCCATCATGGAAGCATGTTGAAACAAGATCCTT 817  
Db 1 CTGAGAGATATCCAAAGAACTAAATGCCATCATGGAAGCATGTTGAAACAAGATCCTT 60

QY 818 CATTAGACCATCTGCTATCGAATTTTAAATCCCTTACCTTGATGACGACTACAGA 877  
Db 61 CATTAGACCATCTGCTATCGAATTTTAAATCCCTTACCTTGATGACGACTACAGA 120

QY 878 ACCTAATGTGTAGATATTGAGAAATGACTCTGGAAGACAAAATTTGGATTGTGAGAAGG 937  
Db 121 ACCTAATGTGTAGATATTGAGAAATGACTCTGGAAGACAAAATTTGGATTGTGAGAAGG 180

QY 938 AGGCTGCTCATATTAATTAATGCCATGCAAAAAGGATCCACCTGCAGACTCTGAGGGCAC 997  
Db 181 AGGCTGCTCATATTAATTAATGCCATGCAAAAAGGATCCACCTGCAGACTCTGAGGGCAC 240

QY 998 TGTGAGAAGTACAGAAATACGCAAGAGAAAGGATCGGGCTGAGGAAGCTCCAGG-CG 1056  
Db 241 TGTGAGAAGTACAGAAATACGCAAGAGAAAGGATCGGGCTGAGGAAGCTCCAGGCGC 300

QY 1057 GCTGATGAGAAGCCAGGAAGCTGAAAAGATTGTTGGAAGAAAATATGAAGAAATAGC 1116  
Db 301 GCTGATGAGAAGCCAGGAAGCTGAAAAGATTGTTGGAAGAAAATATGAAGAAATAGC 360

QY 1117 AAACGAATGCAAGATTGAGATCTCGGAATCTTCAGCAGCTGAGTGTTG-ATGTACTCCA 1175  
Db 361 AAACGAATGCAAGATTGAGATCTTCGGAATCTTCAGCAGCTGAGTGTTGATGTACTTCA 420

QY 1176 TGAATAA 1182  
Db 421 TGAATAA 427

RESULT 15  
ABK43786  
ID ABK43786 standard; cDNA; 430 BP.  
XX  
AC ABK43786;  
XX  
DT 05-JUN-2002 (first entry)  
XX  
DE DNA encoding novel central nervous system protein #366.  
XX  
KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;  
KW hyperproliferative disorder; neoplasm; cardiovascular disorder;  
KW cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;  
KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;  
KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;  
KW adenocarcinoma; reproductive system disorder; testicular feminisation;  
KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;  
KW respiratory disorder; renal disorder; kidney failure; blood disorder;  
KW myocardial infarction; wound healing; cell proliferation; skin aging;  
KW food additive; food preservative; gene therapy; gene; ss.  
OS Homo sapiens.  
XX  
FN WO20015318-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001332.  
XX

PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184644P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0199123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226888P.  
PR 23-AUG-2000; 2000US-0227182P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 26-SEP-2000; 2000US-0235485P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235835P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.



PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249264P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249257P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251989P.  
PR 08-DEC-2000; 2000US-0251990P.  
PR 11-DEC-2000; 2000US-0254097P.  
PR 05-JAN-2001; 2001US-0259678P.  
XX

(HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;

XX WFI; 2001-581633/65.

DR P-PSDS; AAU87456.

XX New isolated nucleic acid encoding a protein for diagnosing, preventing,  
PT treating or ameliorating medical conditions and used as food additives or

PT Preservatives.  
XX Claim 1; SEQ ID NO 376; 837pp; English.  
XX  
CC The invention describes an isolated nucleic acid molecule (I) encoding a  
CC novel central nervous system protein. (I) and polypeptides (III) encoded  
CC by (I), are used to treat a medical condition and in diagnosis of a  
CC pathological condition. Disorders which are diagnosed or treated include  
CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative  
CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders  
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,  
CC angiogenesis, nervous system disorders e.g. Alzheimer's disease and  
CC amyotrophic lateral sclerosis, infections caused by bacteria, viruses  
CC e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders  
CC adenocarcinomas and irritable bowel syndrome, reproductive system  
CC disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes  
CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.  
CC leukaemia, disorders involving neovascularisation e.g. malignancies,  
CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.  
CC acute kidney failure and blood related disorders e.g. myocardial  
CC infarction. The polypeptides can also be used to aid wound healing and  
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to  
CC maintain organs before transplant, for supporting cell culture of  
CC primary tissues, to regenerate tissues and in chemotaxis. The  
CC polypeptides can also be used as a food additive or preservative to  
CC increase or decrease storage capabilities, fat content, lipid, protein,

Query Match 27.3%; Score 395.4; DB 4; Length 430;

Best Local Similarity 97.9%; Pred. No. 4.2e-105;

Matches 418; Conservative 1; Mismatches 6; Indels 2; Gaps 2;

QY 758 CTGAGAGATATCCAAAGAACTAAATGCCATCATCGGAAGCATGTTGAACAGATCCTT 817  
Db 1 CTGAGAGATATCCAAAGAACTAAATGCCATCATCGGAAGCATGTTGAACAGATCCTT 60  
QY 818 CATTAAAGACCATCTGCTATCGAAATTTTAAAAATCCCTTACCTTCATGAGCAGCTACAGA 877  
Db 61 CATTAAAGACCATCTGCTATCGAAATTTTAAAAATCCCTTACCTTCATGAGCAGCTACAGA 120  
QY 878 ACCTAATGTGTAGATATTCAGAAATGACTCTGGAAGACAAATAATTGGATTGTCAAGG 937  
Db 121 ACCTAATGTGTAGATATTCAGAAATGACTCTGGAAGACAAATAATTGGATTGTCAAGG 180  
QY 938 AGGCTGCTCATATAATTAATGCGATGCAAAAAGGATCCACTGCAGACTCTGAGGGCAC 997  
Db 181 AGGCTGCTCATATAATTAATGCGATGCAAAAAGGATCCACTGCAGACTCTGAGGGCAC 240  
QY 998 TGTCAAGATACAGAAAATGAGCCCAAGAGAAAGATCGCGCTGAGGAAGTCCAGG-CG 1056  
Db 241 TGTCAAGATACAGAAAATGAGCCCAAGAGAAAGATCGCGCTGAGGAAGTCCAGGGCG 300  
QY 1057 GCTGATGAGAAAGCCAGGAGCTGAAAAGATTGTTGGAAGAAAATATGAAGAAATAGC 1116  
Db 301 GCTGATGAGAAAGCCAGGAGCTGAAAAGATTGTTGGAAGAAAATATGAAGAAATAGC 360  
QY 1117 AAACGAATGCAAGAAATTCAGATCTCGGAACCTTCAGCAGCTGAGTGTG-ATGTACTCCA 1175  
Db 361 AAACGAATGCAAGAAATTCAGATCTCGGAACCTTCAGCAGCTGAGTGTG-ATGTACTCCA 420  
QY 1176 TGAAGAAA 1182  
Db 421 TGAAGAAA 427

Search completed: October 2, 2004, 22:12:24

Job time : 623 secs

***This Page Blank (uspto)***